

q_i to state q_j with exactly m steps (i.e., with m inputs) and 0 otherwise. Also, the $i-j$ entry of B_m is 1 if it is possible to proceed from state q_i to state q_j with at most m inputs, and 0 otherwise.

We wish to show that for some m every entry in the first row of B_m is a 1, i.e., $b_{0j}^{(m)} = 1$; $0 \leq j \leq \alpha - 1$. Thus, with a suitable input sequence of at most m bits, you can go from state q_0 (the initial state) to every other state.

Upon computing the matrix A_m , we find that for $0 \leq k \leq m\alpha - 1$

$$a_{ij}^{(m)} = 1 \text{ when } \begin{cases} i = k \pmod{\alpha} \\ j = \left\lfloor \frac{k}{2^m} \right\rfloor \end{cases}$$

and

$$a_{ij}^{(m)} = 0 \text{ otherwise.}$$

fig. 5 shows A_2 and A_3 for $\alpha = 6$.

(a)	(b)
$\begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$

Fig. 5—Transition matrices for $\alpha = 6$. (a) A_2 . (b) A_3 .

Observe that the number of 1's in each column of A_m is 2^m for $2^m \leq \alpha$, and that for $2^m \geq \alpha$ the matrices A_m and B_m have all their entries equal to 1. Hence the number of input bits required to go from any state to any other state is no larger than the least integer greater than or equal to $\log_2 \alpha$. Thus, starting in state q_0 , each of the α states can be reached. This shows that α states are both necessary and sufficient for a Turing machine of our form to perform multiplication by α .

JOEL M. WINETT
M.I.T. Lincoln Lab.
Lexington, Mass.

Improved Memory Matrices for the n -Tuple Pattern Recognition Method*

SUMMARY

It is shown that a previous version of the n -tuple pattern recognition method can be made more effective by making certain changes in the learning phase. A means of further increasing readability through judicious choice of n -tuples is described.

INTRODUCTION

Two phases, "learning," and "reading," are used in the n -tuple pattern recognition method.¹ The memory matrix stores information obtained from sample patterns dur-

ing the learning phase, and this stored information is used in the reading phase to identify patterns not previously presented. Obviously, the quality of information stored and the methods' success are directly related, although other factors must also be taken into account.

A storage address for each possible state ($0, 1, 2, \dots, 2^n - 1$) of each n -tuple for each character type (e.g., A, B, C, ...) is provided in the memory matrix. Only 0's and 1's were stored in these addresses in the original version.² A 1 was stored in a given address if the n -tuple in question was placed in the specified state by any one of the given character images presented during learning. Even in this original paper it was suggested that the memory matrix might be used more effectively if frequencies, rather than merely 0's and 1's were stored.³ This suggestion, however, was not exploited at that time. This paper seeks to demonstrate that readability with the n -tuple method is much improved if memory matrices are chosen which are more nearly optimum than are 0, 1 matrices. Variations resulting from value changes of the parameter n are also given.

TABLE I
PER CENT RECOGNIZED CORRECTLY
COMPARISON OF METHODS

Row	Method	40 Sets Read (Same as learned)			10 Sets Read (Different than learned)		
		$n=1$	2	6	$n=1$	2	6
1	0, 1 Matrix	%	%	%	%	%	%
2	Probability Matrix	45.0	48.5	62.0	25	25	24
3	Normalized Within State	70.0	70.0		47	47	
4	Highleyman-Kamentsky Method	77.5	78.5	81.5	58	58	53
5	Highleyman-Kamentsky Method, Zero State Suppressed	76.7	77.2	80.7	59	62	51
6	Maximum Likelihood Method	79.7	83.2	91.7	61	63	61

SOURCE OF DATA

The images used to obtain data given in Table I consisted of 50 sets of handwritten numbers, 0 through 9, for a total of 500 separate images. These numerals were written by 50 different persons at Bell Telephone Laboratories.⁴ Subsequently, the images were individually digitized on a 12 X 12 grid. In all trials reported here, 40 sets were learned; following which, all 50 sets were read. Before learning or reading, each image was centered by placing its center of gravity at a predetermined point on the "retina."

Percentage of characters successfully recognized under a variety of conditions is summarized in Table I. The various methods used were listed in the first column. The second column shows percentage successfully read via each method, using the 40 alphabets already learned. The third column shows percentage successfully read of the 10 alphabets not previously learned. (Ties in all cases are counted as failures.) In each case, results are shown for $n=1$, $n=2$, and $n=6$.

* Bledsoe and Browning, *ibid.*, pp. 225-227.

² Bledsoe and Browning, *op. cit.*, see "Probability," p. 231.

³ Provided to the authors by W. H. Highleyman and referred to by Highleyman and L. A. Kamentsky in "Comments on a character recognition method of Bledsoe and Browning," IRE TRANS. ON ELECTRONIC COMPUTERS (Correspondence), vol. EC-9, p. 263; June, 1960.

used the method quite successfully, on this same data, for the case of $n=1$.⁶ As reported in row 4, the matrix entries were normalized "within character class." Each entry was divided by the square root of the sum of the squares of entries referring to the particular character class. The matrix is thus reduced to a set of unit vectors.

A variation of the Highleyman-Kamentsky normalization technique, in which the zero states of the n -tuples are suppressed, is represented in row 5 of Table I. This corresponds, in the case of $n=1$, to handling "black" image areas only.

THE MAXIMUM LIKELIHOOD TECHNIQUE

A procedure⁷ based on a maximum likelihood technique was employed to obtain the results given in row 6 of Table I. This procedure, for $n \geq 2$, is a generalization of the technique described in detail by Minsky in his "Steps Toward Artificial Intelligence."⁸ Here each entry in the frequency memory matrix is replaced by its logarithm,

⁶ W. H. Highleyman, "An analog method for character recognition," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-10, pp. 502-512; September, 1961. (See also Highleyman and Kamentsky.)

⁷ Other centering techniques which further increase readability were also introduced by Highleyman and Kamentsky.

⁸ W. Harkness, D. T. Laird, and L. L. Pryor, Pennsylvania State University (private communication).

⁹ M. L. Minsky, "Steps toward artificial intelligence," Proc. IRE, vol. 49, pp. 14-15; January, 1961.

* Received July 2, 1961.
¹ W. W. Bledsoe and I. Browning, "Pattern recognition and reading by machine," Proc. Eastern Joint Computer Conf., Boston, Mass., pp. 225-232; 1959.

with zero replaced by some preselected negative number. As Table I shows, "per cent read" improved dramatically when this approach is used, increasing from a rate of 13 per cent to a rate of 83 per cent in the case of $n=2$. (Ties are counted as incorrect.) The maximum likelihood learning procedure is the most efficient yet reported for the n -tuple method.

In regard to all the above methods, in so far as they utilize conditional probabilities, decision functions, and correlation techniques, reference should be made to the work of Chow.¹

SOME OBSERVATIONS

Note that $n=2$ definitely provides better readability than $n=1$, although not by a large amount. It can also be seen that $n=6$ is more impressive for the 40 learned image sets than for the 10 unlearned sets. Readability for $n=6$ on unlearned images probably would improve if learning experi-

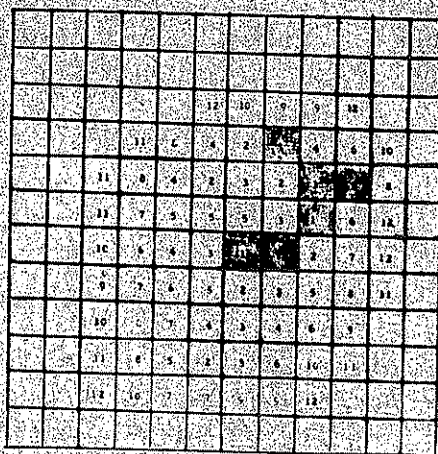


Fig. 1—The 12X12 retina, showing the elements of each of the twelve specially chosen 6-tuples which were used for the results given in Table II. For example, the elements of the first 6-tuple are shown as dark squares.

TABLE II
COMPARISON OF PER CENT READ CORRECTLY FOR RANDOM n -TUPLE AND THE SPECIALLY CHOSEN 6-TUPLES OF FIG. 1

Row	Method	40 Sets Read (Same As Learned)			10 Sets Read (Different Than Learned)		
		$n=1$	$n=2$	$n=6$	$n=1$	$n=2$	$n=6$
1	0, 1 Matrix Method Using Random n -Tuples (From Table I)		13.0	62.0		2	24
2	0, 1 Matrix Method Using the 12 6-Tuples Shown in Fig. 1			98.0			50
3	Maximum Likelihood Method Using Random n -Tuples (From Table I)	79.7	83.2	91.7	61	63	61
4	Maximum Likelihood Method Using the 12 6-Tuples Shown in Fig. 1			99.75			67

ence was substantially increased, say to 40 sets. Generally speaking readability can be expected to increase with n but, at the same time, more learning experience will be required.

Even though the maximum likelihood method scored better than the other methods used in this small study, it would be a mistake to claim that such a result should have been expected beforehand. In fact the method is based on the assumption that the n -tuples are independent. But for most problems in character recognition they are very dependent (especially in the case $n=1$).¹⁰

Results like those reported in Table I inevitably raise an important question: How much more readability can be expected with further improvements in the matrix? Clearly, an optimum set of stored matrix values exists for any given pattern set. Indeed, learning can be described precisely as the attempt to obtain the optimum matrix for specified sets, some methods being superior to others in any given case.

If "number of images correctly read" is accepted as the definition of "readability," readability for a given pattern set can be regarded as a function of many variables, the variables being values recorded in the matrix. In this light, optimization tech-

niques can be employed to seek the optimum matrix for a given set. The result, when found, could then be recorded as the last or "best" row entry in Table I. The actual matrix which provides optimum performance will depend naturally on the particular n -tuples selected, as well as on other system parameters, including n itself. Optimization, therefore, must at least include: finding the "best" n , finding the "best" set of n -tuples, and finding the "best" corresponding memory matrix.

AN ATTEMPT TO OPTIMIZE n -TUPLES

In the calculations represented by Table I, no effort was made to discover optimum n -tuples. In fact, the n -tuples used were deliberately selected on a random basis. Subsequently, one attempt was made to find a more nearly optimum n -tuple set, as follows:

The probability of a given element being touched by any image presented was established for each retinal element. These elements were then rank-ordered, from most to least probable and placed in groups of six. The first twelve 6-tuples chosen by this method are shown in Fig. 1.

This rather small but apparently powerful set was used in conjunction with the maximum likelihood learning procedure. The result of this attempt is shown in Table II, where it is compared with the corresponding result for the randomly chosen n -tuples. Even better results would be predicted if a larger and still more nearly

optimum n -tuple set were to be applied. For example, the methods of information theory might be applied to choose more efficient n -tuples, with each prospective n -tuple treated as a separate information channel.¹¹

Discovery of the "best" n -tuple sets and the "best" weights for them (matrix entries) seems closely related to the "demon" problem of Selfridge,¹² to the operator problem of Uhr and Vossler,¹³ and to the work of Doyle¹⁴ and others.

OTHER REFERENCES

- [1] W. W. Bledsoe, "Further results on the n -tuple pattern recognition method," IRE TRANS. ON ELECTRONIC COMPUTERS (Correspondence), vol. EC-10, p. 96, March, 1961.
- [2] L. Uhr, "A possibly misleading conclusion as to the inferiority of one method for pattern recognition to a second method to which it is guaranteed to be superior," IRE TRANS. ON ELECTRONIC COMPUTERS (Correspondence), vol. EC-10, p. 96, March, 1961.

W. W. BLEDSOE
C. L. BISSON
Panoramic Research, Inc.
Palo Alto, Calif.

¹¹ G. P. Steck, "Statistical model for the Browning-Bledsoe pattern recognition scheme," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-11, pp. 274-282, April, 1962.

¹² O. G. Selfridge, "Pandemonium: A paradigm for learning," in "Mechanization of Thought Process," HMSO, London, pp. 511-535, 1959.

¹³ L. Uhr and C. Vossler, "A pattern recognition program that generates, evaluates, and adjusts its own operators," Proc. Western Joint Computer Conf., Los Angeles, Calif., pp. 550-570, 1961.

¹⁴ W. Doyle, "Recognition of sloppy handwritten characters," Proc. Western Computer Conf., San Francisco, Calif., pp. 133-142, 1960.

Teaching Aid for "Games That Teach the Fundamentals of Computer Operation"

In the above mentioned article¹ Engelbart outlined a method which gives an audience an insight into the mysteries of a digital computer. This method is unique in that a part of the audience is divided into groups where each person acts as a binary element and, with the proper instructions, the operation of a "human-element" computer is demonstrated. In the demonstration the humans are "wired" into a network similar to an actual computer network. In preliminary testing with student groups, the method was very successful. We have added a simple computer to extend the human-simulation "games" and to give the students a transition to the understanding of the physical realization of these basic operations. The simple computer uses relays for all logic operations, where each relay represents one person.

During the demonstration we first proceed to a point where the human element computer is working satisfactorily in a

¹ Received February 27, 1962.

D. C. Engelbart, IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-10, pp. 31-41, March, 1961.

¹ C. K. Chow, "An optimum character recognition method using decision functions," IRE TRANS. ON ELECTRONIC COMPUTERS, vol. EC-6, pp. 247-254, December, 1957.

¹⁰ See Minsky, op. cit., p. 19, last paragraph.

PROPOSAL
FOR
A STUDY TO DETERMINE THE FEASIBILITY OF A
SIMPLIFIED FACE RECOGNITION MACHINE

Submitted To
KING-HURLEY RESEARCH GROUP
422 Washington Building
Washington, D. C.

Submitted By
DR. W. W. BLEDSOE
PANORAMIC RESEARCH, INC.
3946 Fabian Way
Palo Alto, California

30 January 1963

PANDORAMIC RESEARCH, INC.

3946 FABIAN WAY
PALO ALTO, CALIFORNIA
DAVENPORT 1-6592

PROPOSAL

OBJECTIVES

There are presently in use in industry some machines which read preset fixed-font type. Also many experimental machines are reading, some with reasonable reliability, ordinary printed material. A few machines (mainly being simulated on large computers) can read hand-made letters of the alphabet with some success. These machines are based on a number of different principles applicable to pattern recognition problems and while limited success may be claimed for the problem of reading by machine, none of these principles seems powerful enough in present form to recognize a face as belonging to a certain individual out of a large number of individuals.

It is, therefore, proposed to utilize some of the well known methods which have never been applied to facial recognition problems, along with certain lately developed techniques and digital computer simulation, to attempt the solution of a simplified face recognition problem. In this first attempt the work will be restricted to a small number of individuals and to single view high quality pictures. In this way it is hoped that a better understanding of the overall problem can be gained in the minimum time such that the specification of a practical machine might be made. In other words the primary goal of this first phase of the work will be to determine whether presently available methods can be applied to provide a suitable economical solution to the facial recognition problem.

As a matter of opinion, and based upon such considerations as the fact that the unit cost of recognizing a given face among many examples of faces could be a great deal more than could be tolerated for the recognition of a specific letter "A" among all the "A's" in a book for instance; it would appear reasonable to expect that some combination of pattern recognition techniques could be found such that these goals may be met.

METHODS

This research will be based on and, continue from, the pattern recognition research that has been done in this country during the last few years, in particular the n-tuple pattern recognition method first described by B & B in 1959 and since greatly extended. In addition, the investigators knowledge and experience in the areas of (1) optimization of complex systems (2) simulation on digital computers will contribute significantly to this approach.

This work will require the transduction of many pictures of faces of individuals, several for each individual, into digital computer language. The computer, which is simulating a proposed recognition machine, will accumulate information from pictures of known individuals and "learn" to recognize these individuals. The computer is then fed new pictures which it will attempt to identify.

The way in which the information in the pictures is manipulated by machine is, of course, of crucial importance. The methods that might be used for this could utilize the concepts of n-tupling, masking, filters, thresholds, statistical bookkeeping, correlation (comparing vectors), hill climbing, optimization, etc. The investigators have had considerable experience in each of these areas and will use parts of all of them as it appears profitable.

During the next several months an attempt will be made to have a machine give the names of persons whose pictures are presented clearly with a specified view, within a limited number of persons (say 10). Soon one would hope to extend the number of persons to thousands, and eventually to recognize from other than the specific pose.

It is felt that for an eventual recognition machine a certain amount of work by humans may be highly profitable. For example, when giving the machine a picture to be identified the machine operator might be asked to give the sex (if known), race, etc. and to provide scaling information.

Also before the machine attempts an identification it would go through certain normalization procedures such as scaling, standardizing the total amount of light, centering, etc. Also a mathematical technique might be worked out whereby each picture is "rotated" to a frontal view before recognition is attempted. Another possibility is that the machine operator could give the angle of the view and the machine would enter a mode especially designed for pictures of that particular angle.

When the population to be searched becomes very large one could probably never expect to do better than to divide it up into (hopefully) small classes of people who look very much alike. But if these classes are very small then humans could be economically employed to make the final specific identification. One is, after all, confronted by identical twins, so called "doubles" etc., in the normal course of visual recognition and even human beings are sometimes confounded by this problem.

PANORAMIC RESEARCH, INC.

3946 FABIAN WAY
PALO ALTO, CALIFORNIA

DAVENPORT 1-6592

ADDITIONAL INFORMATION

Start with a rectangular picture or image and assume that it is digitized into a grid of black and white squares (Fig. 1).

MASK.

By a "mask" we mean an opaque sheet with a window in it (see Fig. 2) which can be placed over the picture. One then (usually) counts the number of black squares lying in the window (27, using Figs. 1, 2). This gives a measure of the match between the picture and the mask. Certain masks will match A's better than B's and vice versa. By using enough masks and knowing what to look for one can (hopefully) identify an unknown picture by how it fits the different masks. Note that if we have N masks then, for any picture, we obtain N numbers which represent the N measures of match.

A set of N mask can be used in this way to "transform" a picture into a set of N numbers ($m_1, m_2, \dots, m_N$). This vector ($m_1, \dots, m_N$) is thus a new representation of the picture that may be easier to recognize than the original.

The original picture (Fig. 1) can be thought of as a large binary vector. So we transform vectors into vectors by using masks. But in either form these vectors must be recognized or identified.

STATISTICAL BOOKKEEPING

By statistical bookkeeping on known pictures we can keep track of which vectors tended to be A's and which tended to be B's and thereby hopefully identify a new picture as either an "A" or a "B". For example, the simplest method would be to average the vectors (either before or after masking) for each of the category of pictures (average all the B-vectors, etc). In this way we obtain averages $\bar{A}$, $\bar{B}$, $\bar{C}$, etc. Then an unknown image can be compared to each of these averages, and identified as the one it most resembles.

CORRELATION

This comparison is usually done by correlation. If the new unknown vector is $x = (x_1, x_2, \dots, x_N)$ and

$$\bar{B} = (b_1, b_2, \dots, b_N)$$

then the correlation is given by

$$x \cdot \bar{B} = x_1 b_1 + x_2 b_2 + \dots + x_N b_N$$

which is just the ordinary scalar product of two vectors. If x and $\bar{B}$ are a lot alike, then $x \cdot \bar{B}$ tends to be large. Thus one can use the largest of the numbers

$$x \cdot \bar{A}, x \cdot \bar{B}, \text{ etc.}$$

to identify the unknown x .

In actual practice averaging of the original binary vectors (Figure 1) is not too effective.

HILL CLIMBING

If one changes these average vectors, $\bar{A}$, $\bar{B}$, $\bar{C}$, ... in a clever way he will find other vectors, $\bar{A}'$, $\bar{B}'$, ..., which are more effective for identifying the unknown pictures, $\bar{x}$.

The attempt to find better such $\bar{A}'$, $\bar{B}'$, ... involves optimization of a function of many variables (hundreds or thousands) and can represent a very difficult task. Nevertheless such optimization can produce surprising increases in the recognition ability of the procedure.

A harder optimization (or hill-climbing) problem is one of trying to pick the best masks. Out of the fantastically large number of possible masks one tries to choose a few which will be most effective for his type of problem. The mathematical details of how this can be carried out in practice needs more room than I want to take here. But let me say it is still mostly an art and one is never guaranteed complete success.

n-TUPLING

This refers to the concept that Iben Browning and I introduced in our 1959 paper on pattern recognition. By this method we attempt not only to remember whether a particular square is black or white but also its relation to the blackness or whiteness of certain other squares.

Thus, for example, in Figure 1 we note that the three squares marked 1, 2, 3 have the states (white, black, white). If we put "0" for white and "1" for black, then this triple has state (0, 1, 0) for the

particular picture shown. The triple marked i, j, k has state $(0, 0, 1)$. Of course these two triples would take on different states for other pictures.

In practice one uses not just two triples but many (maybe hundreds). Also one could go to higher n -tuples such as 4-tuples, 5-tuples, 6-tuples. In one case we tried 12-tuples. These higher n -tuples are difficult to handle in the computer because the storage required for each n -tuple is of the order of 2^n .

During a "learning" phase the computer will be given a series of known pictures (say a set of "A"'s). The computer will take statistics on the states of the various n -tuples and thereby determine which states are likely when an "A" is being read. A similar procedure is followed while a set of "B"'s are being "learned", then "C"'s, etc. Then during the recognition phase an unknown picture, x , is considered. The various states of its n -tuples, when compared with the statistics that were taken during the learning phase, will (hopefully) determine its identity.

For a given type of picture recognition problem, certain n -tuples will be more discriminating than others. It is hoped that a few powerful n -tuples can be discovered which will be effective for the face recognition problem. One would probably consider moving a particular n -tuple around on the face to see if it "fits" somewhere. Other

complications could be mentioned but the basic idea of an n-tuple, like the basic idea of a mask, will undoubtedly be useful.

THRESHOLDS

Thresholds are sometimes useful for making a process non-linear. For example, after we have counted the number, m , of black squares which show through the window of a mask, we then determine whether m is greater than a given threshold T . If it is we store a "1" corresponding to that mask. In this way a picture is converted, by N masks, to a binary N -vector. One would then adjust T to obtain the best effect.

A similar use of a threshold can be used in all parts of pattern recognition.

FILTER

A filter is any device that converts the given picture into another picture or into a set of numbers. For example, the original transducer which digitizes the picture is a filter. Also a set of masks can act as a filter in that a set of numbers is produced from the original picture. A filter is useful if it converts one representation of a picture into another representation which is somehow easier to identify.

BIBLIOGRAPHY

- Abelson, R. P. and Sermet, V., "Multidimensional Scaling of Facial Expressions," J. Exp. Psychol 63, 546-554, Jan. 1962.
- Abramson, N. and Braverman, D., "Learning to Recognize Patterns in a Random Environment," IRE Tran. Infor. Theory, IT-8, No. 5, 558-563, Sept. 1962.
- Alt, F. F. "Digital Pattern Recognition by Moments," J D C M, 9, 2, 240-258, April 1962.
- Barlow, H. B., "Sensory Mechanisms, the Reduction of Redundancy, and Intelligence," Proc. Sympos. on the Mechanisation of Thought Processes, Nat'l Physical Lab Symp. No. 18, Teddington, Eng., Her Majesty's Stationary Office, London, Vol. 2, (1959), 537-559.
- Beamer, R. E., "A 'Logical Pattern' Recognition Program", IBM J. Res. Develop. 6, 3(July 1962), 353-360.
- Berttram, S., "Application of Hybrid Analog and Digital Techniques in the automatic Map Compilation System," Proc. 1963 Spring Joint Computer Conf., Detroit, Michigan, May - 21-23, 105-112.
- Blackman, S., "Rotational Tomography of the Face", Brit. J. Radio 33, July 1960, 408-410.
- Brain, A. E., Macovski, A., et al, "Graphical Data Processing Research Study and Experimental Investigation," AD 249 873 (PB 155367) November 1960, 28 pages.
- Braverman, David, "Learning Filters for Optimum Pattern Recognition," IRE Trans. IT-8, 4(July 1962), 280-285.
- Braverman, E. M., "Certain Problems in the Design of Machines Which Classify Objects According to an Identifying Feature Which is Not Specified a Priori," Automat. Remote Contr., 21, 10(October 1960), 971-976.
- Braverman, E. M., "Experiments on Training a Machine to Recognize Visual Shapes," Automat i Telemekh (USSR), Vol. 23, No. 3, 1962, 349-364 (Russian).
- Brody, A. L. and Weinstock, S., "Mathematical Theories in Performance, Decision Making, and Learning - A Literature Review," Tech. Doc. Report No. MRL-TDR-62-76, Behavioral Sci. Lab., Wright-Patterson Air Force Base, Ohio, July 1962.
- Buell, D. N., "Chrysler Optical Processing Scanner (COPS); a Character Recognition System Which is Independent of Character Translation, Size or Orientation" Proc. NJCC, Washington, D. C., December 1961, 352-370.

Bertillon, alphonse

II.

- Carroll, K. D., "Character Recognition Devices for Electronic Computers - An Annotated Bibliography," ASTIA SRB-68-11, November 1968.
- Chow, C. K., "An Optimum Character Recognition System Using Decision Functions," IRE Trans. on Elec. Comp., Vol. EC-6, December 1957, 247-254.
- Chow, C. K., "A Recognition Method Using Neighbor Dependence," IRE Trans. EC-11, 5 (October 1962), 683-690.
- Churchill, E., Truett, B., "Metrical Relations Among Dimensions of the Head and Face," WADC Tech. Report 56-621, ASTIA Doc. No. AD-110-629, June 1957.
- Doyle, W., "Operations Useful for Similarity-Invariant Pattern Recognition," ACM Journal, V. 9, 1962, 259-267.
- Doyle, W., "Recognition of Sloppy Hand-Printed Characters," Proc. WJCC (May 1960), 133-142.
- Ehrhardt, S., "Zur Frage der richtigen Beurteilung Morph Ologischer merkmale am Lichtbild von Kopf und Gesicht" (The Problem of Correct Evaluation of Morphological Characteristics of Photography of Head and Face), Acta Geneticae Medicae et Gemellologiae (Roma) 5 (1), January 1956, 104-112.
- Eugen, T. and Levy, M., "Constant Sam Judgments of Facial Expressions," J. Exp. Psych., 51 (6), June 1956, 396-398.
- Eugen, T., Levy, M. and Schlossberg, M., "The Dimensional Analysis of a New Series of Facial Expressions," J. of Exp. Psych. 55 (5), May 1958, 454-458.
- Fayn, V. E., "Automation of the Input of Various Data Types into a Computer," Automatika i Telemekhanika 22, 4 (1961), 535-538, (Russian).
- Firschel, O. and Fischler, M., "Automatic Subclass Determination for Pattern-Recognition Applications," IEEE Trans. Elec. Comp., April 1963, 137-141.
- Frijda, N. H. and Philipsson, E., "Dimensions of Recognition of Expression," J. Abnorm. Soc. Psychol, 66(1), January 1963, 48-51.
- Gaston, G. A., "A Simple Test for Linear Separability," IEEE Trans. Elec. Comp., April 1963, p. 134.
- Gioffre, L., "The Physiognomic Science," Ann Ital Chir. 39, Oct., 1963, 1023-34 (Italian).

III.

- Graham, R. E. and Kelley, J. L. J., "A Computer Simulation Chain for Research on Picture Coding," 1958 IRE WESCON Convention Record, Part 4, 41-46.
- Grullano, V. E., Jones, P. E., Kimball, G. E., Mayer, R. F. and Stein, B. A., "Automatic Pattern Recognition by a Gestalt Method," Inf. Centr. 4, 4(Dec. 1961), 332-345.
- Gwynne-Evans, E., "Pictorial Analysis; So-Called 'Adenoid' Facies," Postgrad Med. V. 19, March 1966, 253-255.
- Hertl, M., "Anatomical and Physiological Principles of Facial Expression," Hippokrates, 36, 15 March 1963, 167-173 (German).
- Hawkins, J. R., "Self-Organizing Systems - A Review and Commentary," Proc. IRE, Jan. 1961, pp. 31-48.
- Hayward, H. L. and Milton, A. M., "Pattern Recognition for Diagnosis of Facial Characters," Data Processing for Scientific Engineering, Sept. 1963, 31-34.
- Hightleyman, W. H., "An Analog Method for Character Recognition," IRE Trans. EC-10, 3 (Sept. 1961), 502-512.
- Hightleyman, W. H., "Linear Decision Functions with Applications to Pattern Recognition," Proc. IRE, Vol. 50 (6), June, 1962, 1501-1514.
- Hoare, C. A. R., "Quicksort," Comput. J. 5 (1962), 10-15.
- Hoff, M. E., Sr., "Learning Phenomena in Networks of Adaptive Circuits," TR Report No. 1554-1, Stanford Electronics Lab., July 1962.
- Hu, M.-K., "Visual Pattern Recognition by Moment Invariants," IRE Trans. IT-8, 2(Feb. 1962), 179-187.
- Innes, Daphne J., "Filter -- A Topographical Pattern-Separation Computer Program," Proc. EJCC, New York, New York, Dec. 1965, 25-37.
- Ishizawa, T., "A New Process of Character Recognition Using a Sieving Method," Prof. Sp. on Automata and Automatic Control of the Inst. of Elec. Comm. Engrs. of Japan, Fujinicho, Chiejodaka, Japan (June 23, 1961) Ab. IRE Trans. IT 28, 1 (Jan. 1962), 67.
- Jakowatz, C. V., Brennan, E. J. and Geyer, J. M., "One-Dimensional Waveform Recognition Research," Report 62-EC-6225E, General Electric Company, December 1962.
- Joet, G., "Research Based on Artistic Anatomy of the Criteria Designed to Guide the Surgeon in Plastic Operations on the Face," Ann Chir. Plast. 7 (3), March 1962, 39-46 (French).

IV.

Julesz, B., "Toward the Automation of Binocular Depth Perception," Proceedings of the IFIP Congress, Munich, (1962).

Julesz, B., "Visual Pattern Discrimination", IRE Trans. on Info. Theory, IT-8, No. 2, 84-92.

Kamentzky, L. A., "Pattern and Character Recognition Systems -- Picture Processing by Nets of Neuron Like Elements," Proc WJCC San Francisco, Calif., Mar. 3-5, 1959, 304-309.

Kasner, H. and Steinbuck, H., "Adaptive Systems in Pattern Recognition," Proc. IEEE Trans. Elec. Comp., Dec. 1963, pp. 822-834.

Karl, E. R., "Graphic Image Processor," IBM Tech Disclosure Bull., 5 Oct. 1962, 54-56.

Killian, H., "Medical Physiognomy," Medizinische, Jan. 7, 1956, 30, 45-71.

Kirsch, R. A., Ray, L. C., Cahn, L. and Urban, G. H., "Experiments in Processing Pictorial Information with a Digital Computer," Proc. EICC, Wash., D.C., Dec. 9-13, 1957, 221-230.

Kovacs, L. S. G. and Joseph, H. M., "Image Processing," IRE Proc., 43, No. 3, May 1958, 560-570 (St. Britain).

Langdon, J., "The Perception of 3-D Solids," Quart. J. Exptl. Psychol., 7, 1955, 19-36.

Levy, H., Schlossberg, H., "Woodworth Scale Values of the Lightfoot Pictures of Facial Expression," J. Exp. Psychol. 60 (5), Aug. 1960, 121-125.

Lewis, P. M., "The Characteristics Selection Problem in Recognition Systems," IRE Trans. IT-8, 2(Feb. 1962), 171-178.

Labbech, J. K., "Self-Optimizing Filters Employing Orthogonal Systems of Functions," J. Electronics Control (1) 13 (1963), 365-383.

Mach, R. E. and Gardner, T. L., "Rectification of Satellite Photography by Digital Techniques," IBM J. Res. & Dev., 6, July 1962, 270-283.

Marill, T. and Green, D. M., "On the Effectiveness of Receptors in Recognition Systems," IEEE Trans. on Info. Theory, IT-9, Jan. 1963, 11-17.

McGermick, B. H., "The Illinois Pattern Recognition Computer - ILLIAC III," IEEE Trans. on Elec. Comp., Dec. 1963, 791-813.

- Merrill, T. and Green, D. M., "On the Effectiveness of Receptors in Recognition Systems," IRE Trans. IT-9, 1(Jan. 1963), 11-17.
- Ming-Kuei Hsi, "Visual Pattern Recognition by Moment Invariants," IRE Trans. on Info. Theory, IT-8, No. 2, Feb. 1962, 179-187.
- Minot, O. N., "Automatic Devices for Recognition of Visual Two-Dimensional Patterns: A Survey of the Field", Report 1959, U. S. Navy Electronics Lab., 13 June 1961.
- Minsky, M., "A Selected Descriptor - Indexed Bibliography to the Literature on Artificial Intelligence," IRE Trans. on Human Factors in Electronics, March 1961, pp. 39-59.
- Minsky, M., "Steps Toward Artificial Intelligence," Pro. IRE, Jan. 1961 8-20.
- Nadler, M., "An Analogue-Digital Character Recognition System," IEEE Trans. Elec. Comp., Dec. 1963, pp. 514-521.
- Nash, H., "Stereotyped Association to Schematic Faces," J. Genet. Psych., 93 (1), Sept. 1958, 149-152.
- Chajnos, M., Stark, L., Whipple, G. and Yasui, S., "Computer Pattern Recognition Techniques: Some Results with Real Electrocardiographic Data," IRE Trans. on Bio-Medical Electronics, VOLUME 10, No. 2 July 1963.
- Petrov, Yer. E., "Two Dimensional Bundles of Geometric Objects in K_1 ," Mat. Sb. N.S. 46 (85) (1955), 291-314 (Russian).
- Pinaky, M. S., "The Extrapolation of Random Vector Processes and the Quantity of Information Contained in a Stationary Random Vector Process Relative to Another one Stationarily Connected with it," Dok. Akad. Nauk. SSSR, 121 (1958), 49-51 (Russian).
- Ridgway, W. C., III, "An Adaptive Logic System with Generalizing Properties," TR No. 1554-1, Stanford Electronics Lab., April 1962.
- Rosenfeld, Aerial, "Automatic Recognition Techniques Applicable to High-Information Pictorial Inputs," IRE Internat. Convention Record, Vol. 10, Pt. 4, 1962, 114-123.
- Ross, Douglas T., "A Generalized Technique for Symbol Manipulation and Numerical Calculation," Comm. ACM (1961), 147-156.
- Scharr, D. R. and Jouis, J. G., "The Photogrammetric Cloud Measurement Program," AD 257 774, Dec. 1965, 47 pages.

VI.

- Sebestyene, George S., "Decision-Making Processes in Pattern Recognition," (ACM Monograph Series) MacMillan Co., New York, 1962, 162 pg.
- Sendrall, M., "Legitimacy of a Science of the Face," Concours Med. 81:3177-3179, 20 November 1959 (French).
- Sherman, H., "A quasi-topological Method for the Recognition of Line Patterns," Proc. Internat. Conf. Inf. Process. UNESCO, Paris, (June 15-20, 1959), 232-236.
- Shibata, M., Hirotsu, T., Tsuruzono, M. and Teranishi, H., "Photographometrical Judgement of Facial Resemblance Between Half Brothers," Jap. J. Leg. Med., pg. 130-135, V. 17, March 1963 (Japanese).
- Singer, J. R., "A Self Organizing Recognition System," Proc. WJCC, Los Angeles, Calif. (May 9-11, 1961) 345-364.
- Stemmer, L., "Facial Kinetographs - Description of a Method for the Study of New Treatments for Facial Paralysis," Ann. of Otolaryngology, Rhinology and Laryngology (St. Louis) 68(1), March 1959, 208-221.
- Stevens, Mary Elizabeth, "Abstract Shape Recognition by Machine," Proc. EICC, Washington, D. C., Dec. 1961, 332-351.
- Strobl, S., "Over het Bieden Naar de Nationaliteit van Personen op Basis Van Hun (Gefotografeerde) Uiterlijk" (Determination of Nationality on the Basis of External (Photographed) Characteristics), Neder Landesch Tijdschrift voor de Psychologie en Haar Gransgebieden (Amsterdam) 11(6), 1956, 528-536.
- Sublette, J. H., and Kalte, J., "Character Recognition by Digital Feature Selection," RCA Rev. 23 (Mar. 1962) 6, UFO.
- Swift, J. D., "Isomorph Rejection in Exhaustive Search Techniques," Proc. Sympos. Appl. Math., Vol 10, (1960), 195-200.
- Uhr, L., "Pattern Perception and Learning Programs," Commun ACM, 3, November 1960, pg. 597.
- Uhr, L. and Vossler, C., "A Pattern Recognition Program that Generates, Evaluates, and Adjusts its own Operators," Proc. WJCC, Los Angeles 19(May 9-11, 1961), 363-369.
- Weike, R. W., "Rotating Raster Character Recognition System," Comm. Elec. 36 (Sept. 1961), 353-359.
- Widrow, B., "An Adaptive Adaline Neuron Using Chemical Memistors," TR No. 1853-2, Stanford Electronics Lab., 17 October 1960.

VII.

Widrow, B. and Hoff, M. E., Jr., "Adaptive Switching Circuits," TR No. 1553-1, Stanford Electronics Lab., June 1960.

Wong, E. and Thomas, J. B., "On the Multidimensional Prediction and Filtering Problem and the Factorization of Spectral Matrices," J. Franklin Inst., 272 (1961), 87-99.

"Machine Perception of Forms by Means of Assessment and Recognition of Gestalts," The Univ. of Mich., Mental Health Res. Inst., Ann Arbor, Preprint No. 34, Oct. 1959.

"Networks for Pattern Perception," Proc. Nat'l Electronics Conf., V. 14 (1959), 187-189.

"Pattern Detection and Recognition," Proc. IRE, 47, Oct. 1959 (1737-1752).

West Midlands Police

PM



Stoney Stanton Road
Police Station
COVENTRY, CV6 5DG
England

~~M DIVISION HEADQUARTERS~~
~~LITTLE PARK STREET~~
~~COVENTRY~~
~~CV1 2JX~~

~~TELEPHONE No. 0203 555 333~~

ALL COMMUNICATIONS SHOULD BE ADDRESSED TO "THE CHIEF SUPERINTENDENT"

YOUR REFERENCE

OUR REFERENCE

DATE

Professor Woodrow W Bledsoe
MCC 9430 Research Boulevard
Building 1
Austin
TEXAS 78759

19 February 1987

Dear Professor Bledsoe

The Chief Constable of this Force has been kind enough to allow me to indulge in a part time Ph D study into the ways in which we can use computers to recognise the faces of criminals recorded by bank cameras, or during riots, etc. etc.

Presently I am registered as a student with Imperial College in London reading under the tutelage of Professor Igor Aleksander, FRS. The main thrust of the study is to use Professor Aleksander's 'Wisard system' which is a development of the early perceptron devices, but in this case using quasi parallel techniques.

I have had a great deal of difficulty in tracking down copies of your earlier work published under the following titles -

The Model Method in Facial Recognition
Report No. PRI - 15 Panoramic Research Inc. Palo Alto.
1964

Man Machine Facial Recognition
Report No. PRI - 22 Panoramic Research Inc. Palo Alto.
1964.

I wonder if I could impose on your good nature and ask for a copy of these reports? I am well aware that after such a long period of time it may be that you have filed them in a library somewhere. If that is the case, may I have details of where I could obtain them and I will do so.

Kindly

R B STARKEY, B Sc. M Sc.
Chief Inspector

PHONE MESSAGE	TO	WLB		DATE	7-31	TIME	AM PM
	FROM	M	Diana Stasiuk	AREA CODE	716	NUMBER	631-3730
	OF	Seipar Inc.					
	MESSAGE	P.O. Box 185 Buffalo, NY 14221					
				SIGNED msg. ctr.			
PHONED	☒	CALL BACK	☒	RETURNED CALL	☐	WANTS TO SEE YOU	☐
						WILL CALL AGAIN	☐
						WAS IN	☐
						URGENT	☐

in f R
u

Attn: Diana Stasiuk
Seipar Inc.
26 W. Spring St.
Williamsville, NY 14221

Please Mail Xerox

to there to

Do not send the original

Thank

WLB



Figure 2. Examples of Photograph Pairs Used in the Study

Save this one too

12 August 1964

PRI:15

THE MODEL METHOD OF FACIAL RECOGNITION

W. W. BLEDSOE

Panoramic Research, Inc., 2670 Hanover Street, Palo Alto, California.

THE MODEL METHOD IN FACIAL RECOGNITION

W. W. BLEDSOE

ABSTRACT

This paper discusses the general problem of facial recognition by machine, introduces and explains a model method of feature location, discusses procedures for feature classification and facial recognition, and finally discusses several man-machine techniques in facial recognition.

12 August 1964

INTRODUCTION

One of the most challenging areas of pattern recognition is the identification of human photographs by machine. In this endeavor the machine is expected to disclose the identity of the person whose photograph it is shown. Since the difficulty of this problem will depend greatly upon the conditions under which it is considered, we will be more specific.

In this paper we will consider only the case where the following conditions are satisfied:

1. There is a large number of people in the population considered (perhaps more than one million).
2. There is exactly one photograph of each person in the file.
3. Each photograph is oriented in an upright position.
4. The pose for each photograph in the file must lie between the frontal and profile. (See Figure 1)
5. Other variables, such as tilt angle, lighting direction and intensity, and facial expression are not fixed.
6. For recognition, a photograph (other than the one in the file) of a person of the population is presented to the machine and it is expected to
 - (a) Identify the person photographed, or
 - (b) Make a small list of names which includes the correct identity.

There are many aspects of this problem which make it especially difficult. Some of them are:

1. There is a large number of categories. Whereas the alphanumeric character recognition problem may have only 36 categories, the facial recognition problem can have millions or even billions.
2. There is a small number of samples available (usually one) for each category (person).
3. The variability is extensive. It includes:
 - (1) Head rotation (from frontal to profile), and tilt.
 - (2) Lighting intensity and angle.
 - (3) Photograph size (scale).
 - (4) Facial expression.
 - (5) Aging.
 - (6) Hair growth.

(See Figure 2)

It is readily apparent that a simple minded approach will not be successful for this problem. For example, an ordinary correlation technique will fail completely to identify a photograph where the head rotation of the new photograph is slightly different from the one of the same person in the file. (See Figure 3). In fact, correlation might choose a different person but with the same head rotation, tilt, and

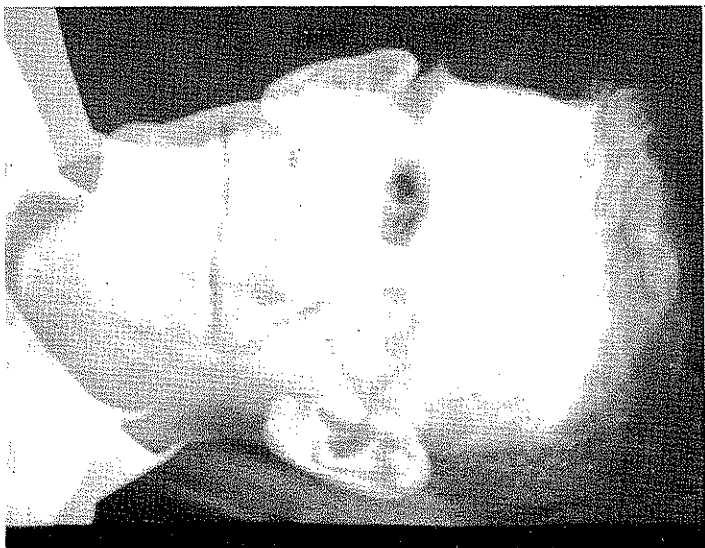
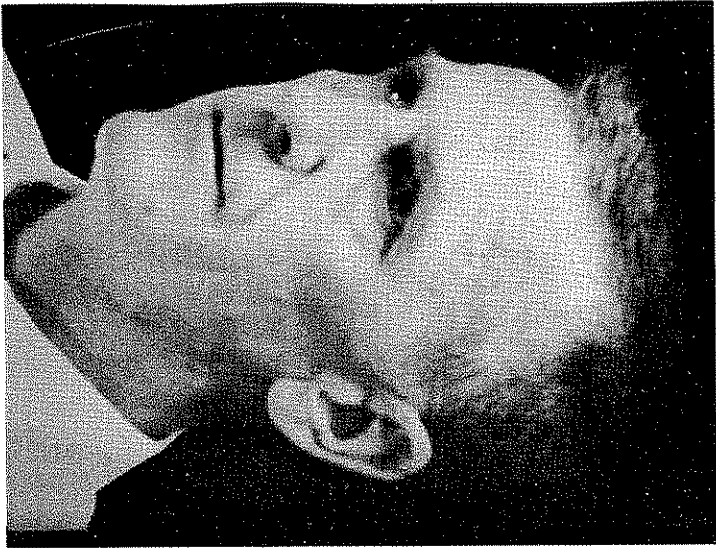


Figure 3

Double exposure shows that the two poses of the same subject have very little in common when considered on a point-by-point basis.

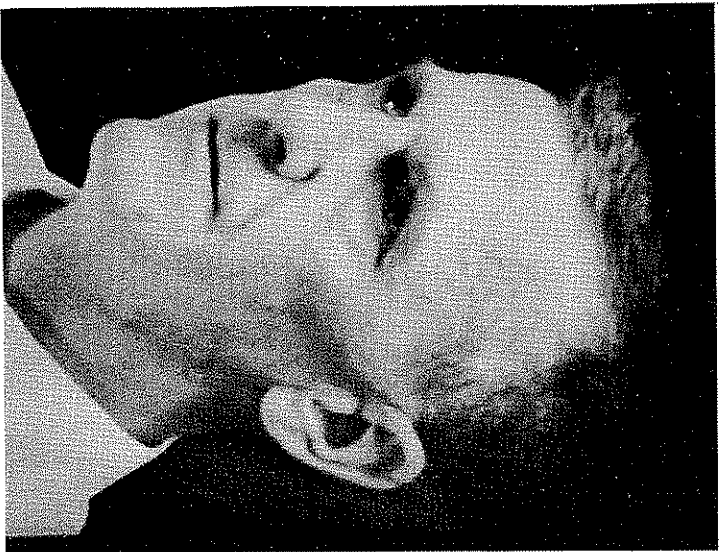


Figure 4

Double exposure shows that two different subjects are surprisingly similar on a point-by-point basis.

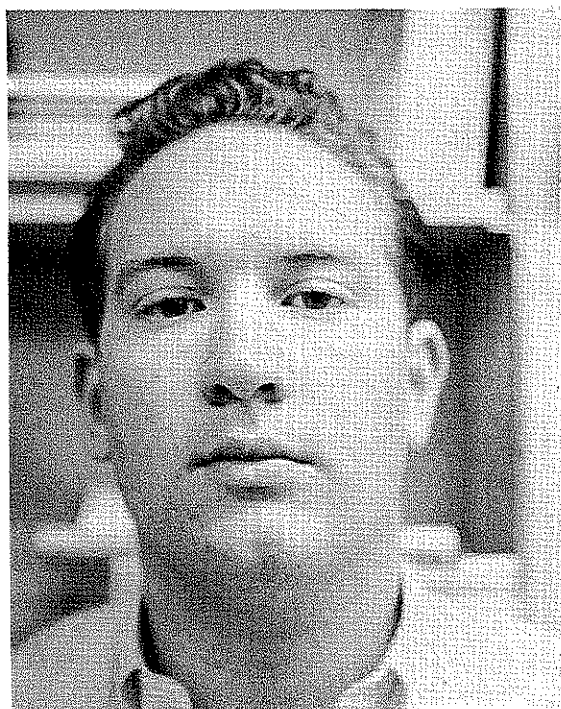
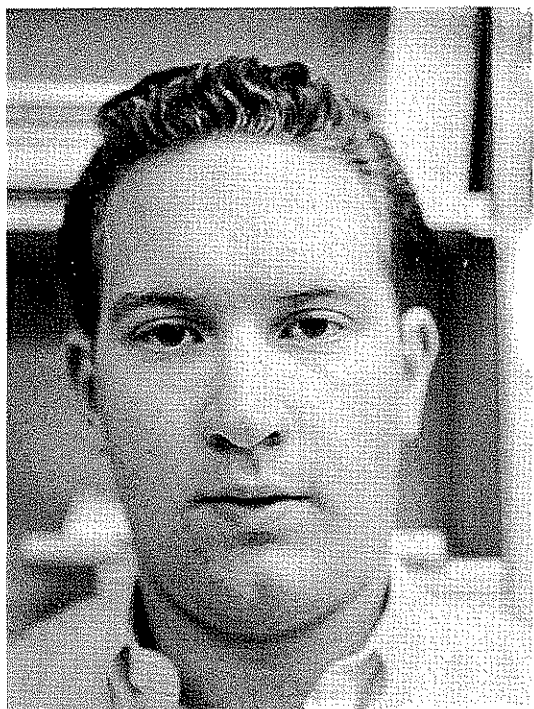


Figure 2a
Variations in Head Tilt

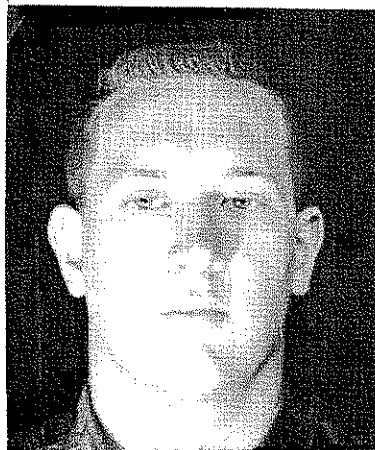
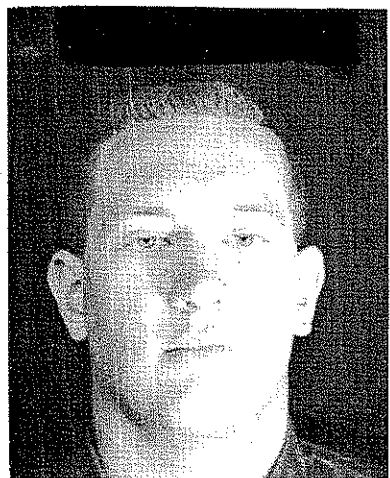


Figure 2b
Variations in Lighting Direction

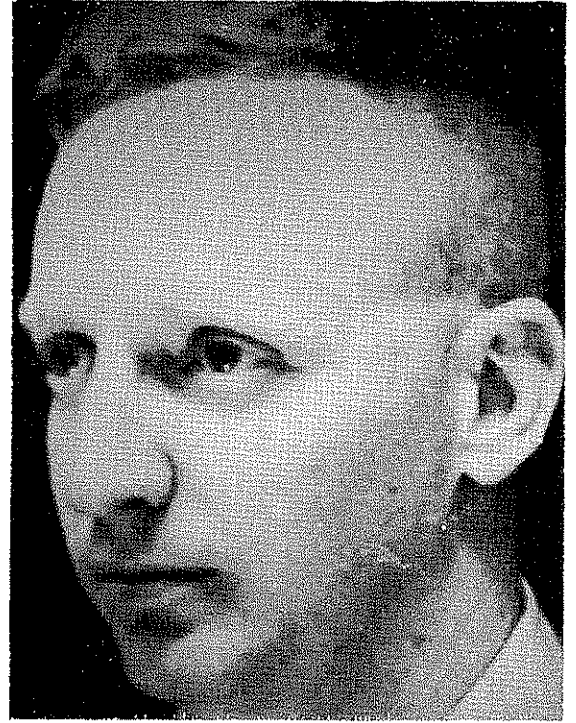


Figure 2c
Variations in Scale and Light Intensity

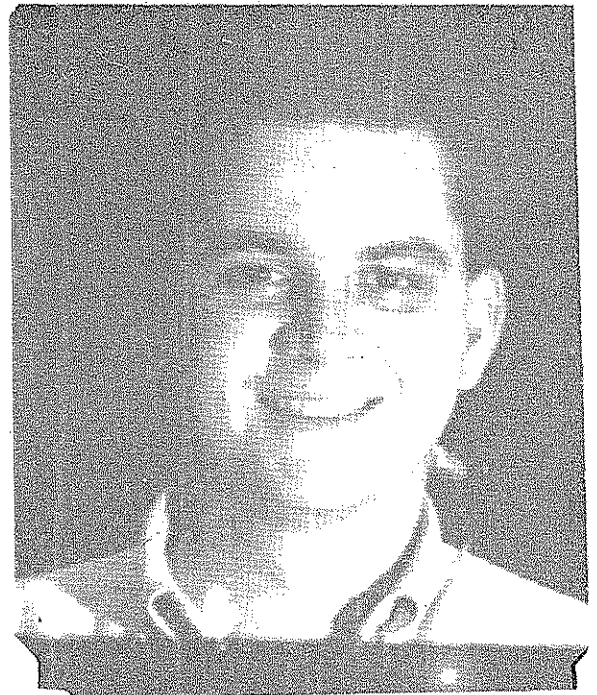
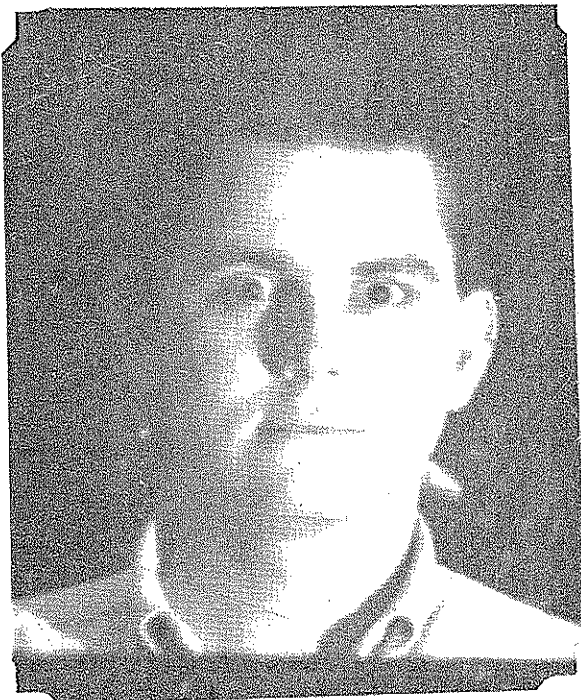


Figure 2d
Variations in Expression

lighting. (See Figure 4) Similarly, failure is predicted for any method that does not employ a measuring system which is invariant under the changes listed under 3 above.

THE FEATURE CLASSIFICATION APPROACH.

It has been proposed¹ that the machine should "find" and "classify" each of a series of features of the face and use this set of classifications to identify the photograph. That shall be our general approach here.

But how can a machine be made to locate a certain feature such as the nose? At first blush this would appear to be a two-dimensional searching problem, but it is far more complex than that. For example, a simple change of lighting can completely displace the supposed location of a feature. Even so, such a search can be carried out successfully if coordinated with other necessary activities, such as determining the lighting angle, head rotation, etc.

Our feeling is that the location of the principal features, such as eyes, nose, ears, mouth, etc. and the determination of the principal parameters of photography, such as head rotation and tilt, lighting angle and intensity, scale, etc., comprise a major part of the problem. In fact, we feel that a feature cannot be satisfactorily located until its principal lines and parts have been completely located and traced. For

1. Bledsoe, W. W., et. al., "Facial Recognition Project Report," Panoramic Research, Inc., Publication PR-7, 6 March 1964.

if we are to make measurements from the corner of an eye, for example, we must know exactly where that corner is. In this sense a feature is essentially identified when it is properly located and the photography parameters are known and can thereafter be easily classified.

In order to search for a feature it becomes necessary to know its likely location and likely shape. But its location and shape will depend greatly on the rotation and lighting angle. On the other hand the rotation and lighting angle cannot be properly determined until we have successfully located some features. It would seem then that neither task can be performed first because it needs the result of the other. Our approach will be to continually use our best guess of the photography parameters in our search for features and, as we gain more information about the features, to continually update and correct our photography parameters.

To carry out such a complicated feedback process, even in a digital computer, requires a large and intricate program unless the problem can be made more tractable. It is desirable to organize the procedure in such a way that the successive steps in the search are similar in nature, so that the whole process is an iteration of almost identical parts.

COMPUTER AND SCANNER.

In what follows we shall tacitly assume that a digital computer is the machine we have in mind for carrying out the prescribed process,

although no such restriction is necessary. It is assumed that the light reflectivity information has been fed to the computer from an optical scanner for a sufficiently dense grid of points on the photograph.

ORIENTATION.

This writer believes that "orientation" is of prime importance in recognition. We first get ourselves familiar with the general surroundings and then proceed to recognize the detail parts. A similar thing happens in facial recognition. The observer first recognizes the object in the photograph as the head of a person, locates the general position of the head and its principal parts and then begins recognition. This process is so fast that it may seem instantaneous and in fact the two steps may proceed in very close succession; but, no actual recognition is possible without the first orientation step. In fact, there probably takes place a whole series of steps of orientation and partial recognition which extend from a vague awareness that "this is a human head," through the determination that "it is a partial profile of a male," to the recognition that "it is John Smith with hat on."

HYPOTHESIS GENERATION

We believe that as these steps in recognition unfold, the brain continually generates hypothesis as to what to look for next to accomplish further recognition. Feedback is rapid and important; as the previous hypothesis is verified stronger evidence is thereby provided for a further hypothesis and so on until recognition is complete.

It is not our purpose here to argue for a theory of human recognition, but rather to set down a reasonable procedure that can be followed by recognition machines. Thus in what follows we will outline a procedure which causes the machine to continually hypothocate what the object face is like and to continually check this hypothesis, and through feedback, update and regenerate this thpothesis.²

THE MODEL APPROACH.

We believe that this approach can best be implemented by the use of a three-dimensional model, which we allow to vary as we proceed through the location and identification of features. The projection of this model provides a two-dimensional map which acts as a guide in the search. (See Figure 5) The photograph, which we are trying to identify, is essentially being compared with the map, but not by a point by point correlation. The (3-dimensional) model will start as a "standard" head with a given set of photography parameters. These parameters and the model's shape will gradually change to compensate for discrepancies found between the map and the photograph.

This three-dimensional model can either be a real physical model or one which is simulated in the computer. The present state of the art of varying physical models precludes the first at this time and leaves the simulation as the only practical solution.

2. This has a similarity to the Problem Solving Programs of Newell, Simon and Shaw. See Reference 3.

In order that the necessary changes in the model be made in the most natural and likely-to-succeed way, the features are numbered and searched for in a certain predetermined way. Thus, for example, the general location of the head is determined several steps before the eye locations are attempted. Each feature is searched for in the position and shape predicted by the map, and as features are found the model is corrected to make the map better "fit" the photograph.

The changes in the model will tend to be those which least distort the model and, where possible, which can be made by changes in its photographic parameters. Thus, for example, if both the left and right sides of the head are found displaced a distance, x , to the left of the corresponding positions on the map then the model is translated a distance, x , to the left; but if the left side is displaced left and the right side is displaced right then an increase in scale for the model may be in order. Similarly, after the head is located, if both eyes are displaced horizontally left this might be corrected by a rotation of the head, or if displaced vertically a tilt of the model might be in order.

LIGHTING.

Lighting is very important in processing photographs because the shadows play a major role in making lines on the face. Thus it is of utmost importance that light intensity and (especially) direction be determined correctly.

Light intensity can be handled fairly well by normalization of the total light reflected from the face, but direction is a more formidable problem.

After the general outline of the head has been determined, a first approximation of the lighting direction can be obtained by comparing the total light reflected from each of the two sides of the face. (See Figure 6) This initial estimate serves to help predict the location of shadows from the nose, etc., and the actual location of these shadows, in turn, can be used to better estimate the lighting direction.

AN APPARENT CONTRADICTION.

Strictly speaking, no three-dimensional information can be deduced absolutely from a single two-dimensional photograph. (The object being photographed might have itself been a two-dimensional photograph!) But then why can people routinely use such photographs for identification and other three dimensional purposes? And why can a one-eyed person function quite well without stereoscopic vision? The answer lies, we believe, in laws of probability. It is not likely that we are looking at a man whose face is completely flat but is colored so that it looks three-dimensional. By experience we learn what to expect and what not to expect. We allow only certain variations about the norms. Shadows help immensely because they give information about the relative heights of objects. A slight rotation of the head gives some information on the relative heights of the parts of the nose, given that we know the general shape of the nose.

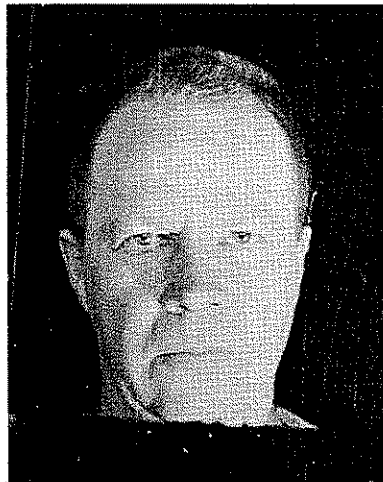


Figure 6

Total reflectivity on the right side of the face is about six times that on the left side, indicating a light angle of about 45° from the right.

This then is the basis upon which the model approach can work. The model is changed from the most general and most likely to the specific in such a way that the changes made are the most probable way to attain the desired results.

FEATURE NUMBERING.

As mentioned above the features on the map are numbered so that they can be searched out in a predetermined order. Certain points on the features of the model are designated as key points and play a role in locating the features on the photograph. These key points are projected onto the map (See Figure 7) and are kept track of during the rotations and changes in the model. These key points on the map then serve as starting points for locating features on the photograph. Thus at all times the computer has the location of these key points so that it does not have to search for the features on the map as well as the photograph. But once the general position of the feature is located it is traced and compared with the corresponding feature information on the map, any discrepancies are detected and corrected by changes in the model which are reflected as changes in the map. This changing map then becomes, in effect, part of the computer program for searching the photograph, because it provides, on a changing basis, information on where to search and for what to search. In this way the general computer program can be greatly simplified.

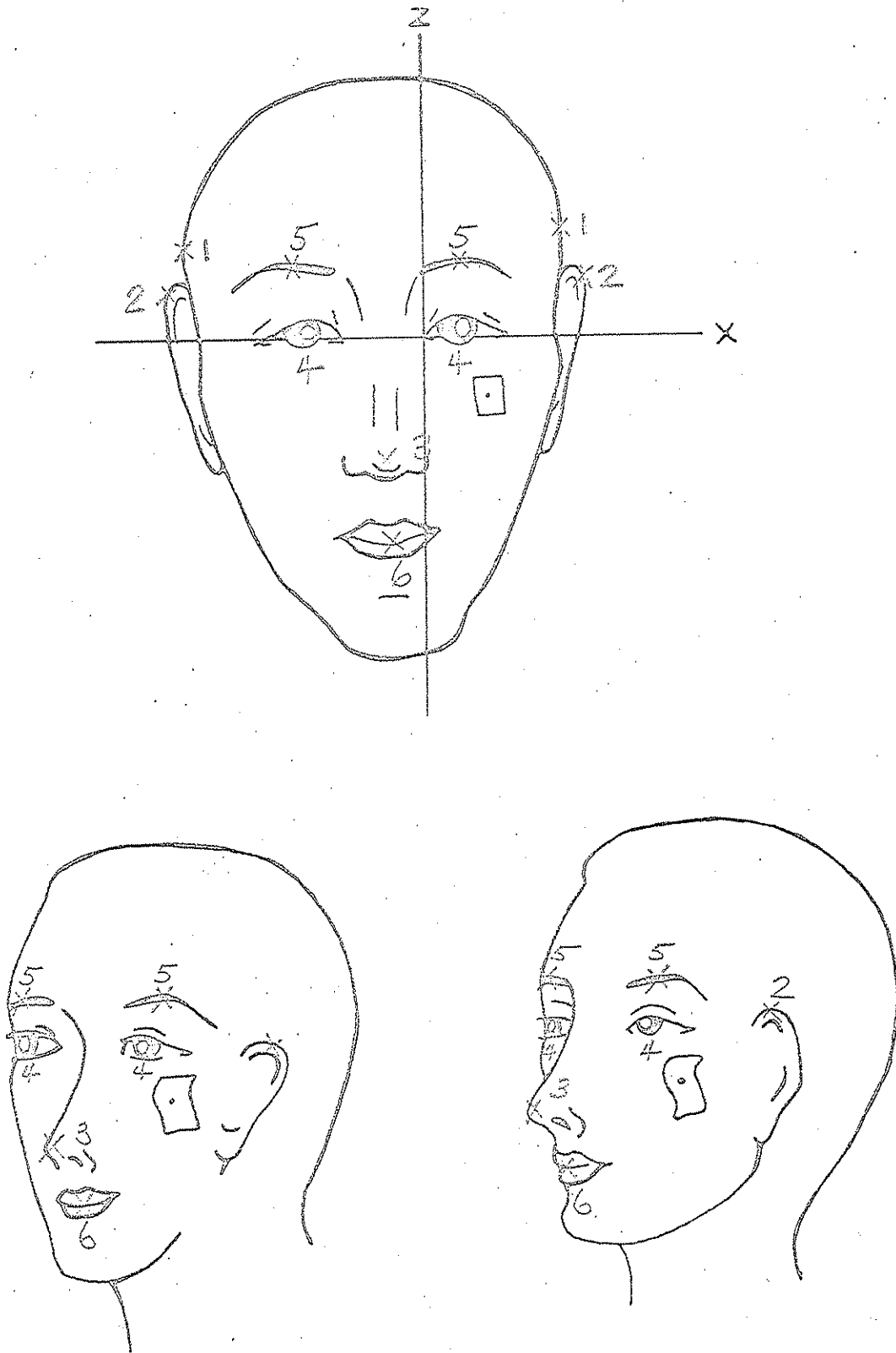


Figure 7

Key points for ordered search

FEATURE CLASSIFICATION.

Once the features of the face have been found, each localized area about a feature can be treated as a recognition problem in itself and the feature classified in a routine manner³. (See Figure 8) Then the totality of these classifications can be utilized to identify the whole face.

Another more direct approach might be attempted for the classification of features. Instead of locating the feature, as outlined in the preceding pages, we might forego their location altogether and rely only on the general location of the head, and classify features directly from the whole photograph. Thus the machine might be asked to determine whether the person in the photograph has a long, medium, or short nose, and similar questions for other features. For this purpose, as well as for classifying the localized features, a straight forward classification procedure⁴ could be employed. It is expected, however, that unless the feature is isolated the classification of it will be difficult indeed and the situation be highly complicated by the uncertainty of the location and shape of the feature and the presence of the other features of the face. By locating the feature first, the variability is greatly reduced and correspondingly the recognition device required is markedly improved

3. See, for example, Appendix V, page 5, of Reference 1.

4. See, for example, references 6-9.

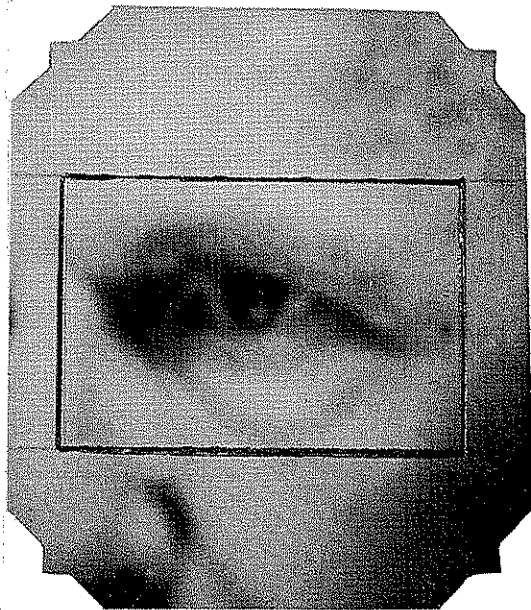


Figure 8

An eye which has been localized for
feature classification.

with respect to size and speed.

There is also a more fundamental objection to this procedure. It is believed that any partition which will classify noses (say) on all types of complete faces, will of necessity at the same time also classify the complete faces; thus, if this is the case, there would be no point in feature classification. Hence we prefer to localize the features first.

FEATURE MEASUREMENTS.

In addition to, or instead of, classifications of the features that have been located we can also make measurements of, and between, the features.⁵ For example, we can measure the distance between the eyes (and properly correct this distance to compensate for facial rotation), and the distance between the eyes and mouth, and the height of the ears, etc. These measurements can be used, with or without feature classifications, to help classify the photographs.

MAN-MACHINE INTERACTION.

It may become desirable, especially in some earlier stages of an application, to lessen the problems imposed on the computer by a certain amount of cooperation between man and machine. The following is a list of suggestions along this line:

5. H. C. McDonald and H. W. Rogers, The Classification of Police Photographs, DeVoss and Co., Los Angeles, Calif., 1941.

A. Input parameters.

The (human) operator may be called upon to give the machine certain parameters at the time the photograph is presented to the scanner.

- | | |
|--------------------|-----------|
| 1. Light intensity | 5. Age |
| 2. Light direction | 6. Weight |
| 3. Head rotation | 7. Race |
| 4. Head tilt | 8. Sex |

Some of these parameters, such as the first five, can be estimated by comparing the photograph to a set of standards (See Figure 9).

B. Location of features.

By special marks, perhaps using a series of different colors, certain features can be located on the photograph by the operators.

Thus a dot could be placed at -

1. Tip of the nose
2. Corner of each eye
3. Corners of the mouth
4. Extent of the ears
5. Point of the chin
6. The middle and ends of the hairline
7. Etc.

These positions also could be given by a light pencil on a display (CRT) tube, on which the photograph is projected, on request from the computer.

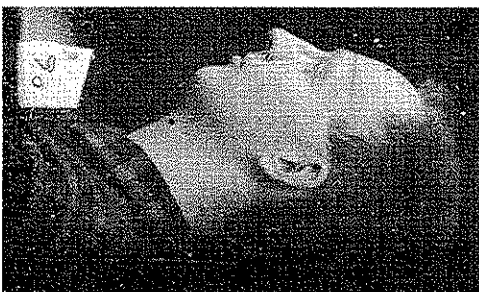
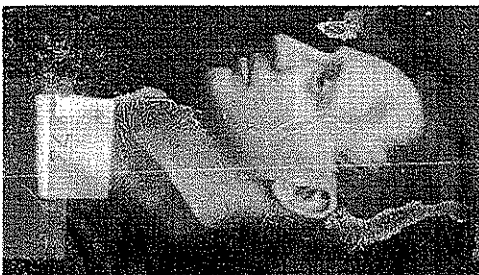
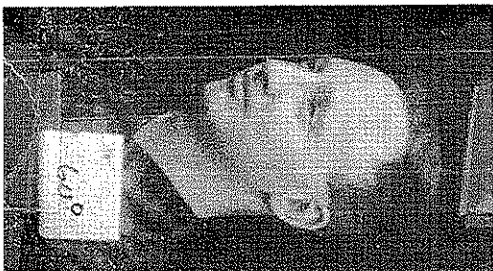
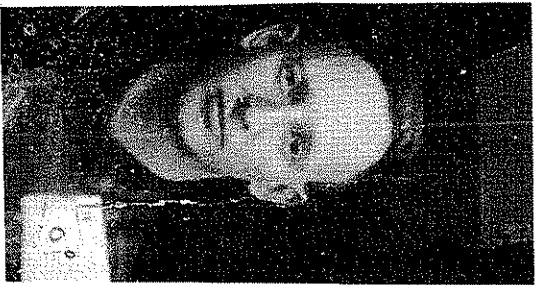


Figure 9

A series of poses that can be used by an operator to estimate rotation angle.

Then the computer can use these points to calculate distances of the type described above under feature measurements.

C. Feedback.

In an application where features are located by machine, the computer's location of the various features can be verified by the human operator through the use of a display tube. The computer displays the original photograph on its display table, indicating by appropriate marks its location of the features, and also giving the photographic parameters it has determined. If these don't satisfy the operator he can make appropriate changes.

REFERENCES

1. W. W. Bledsoe, et. al., "Facial Recognition Project Report," Panoramic Research, Inc., Publication PR-10, 6 March 1964.
2. See list of references at the end of reference 1.
3. Report on a General Problem - Solving Program, " The RAND Corp., P-1584, 9 February 1959.
4. H. C. McDonald and H. W. Rogers, The Classification of Police Photographs, DeVoss and Co., Los Angeles, California, 1941.
5. H. C. McDonald, "Identi-Kit Manual, A System of Modern Visual Identification," Identi-Kit Div., Townsend Co., Santa Ana, Calif.
6. Highleyman, W. H., "Linear Decision Functions with Applications to Pattern Recognition," Proc. IRE, Vol. 50 (6), June 1962, 1501-1514.
7. Widrow, B., and Hoff, M. E., "Adaptive Switching Circuits," TR No. 1553-1, Stanford Electronics Lab., 17 October 1960.
8. Alt, F. F., "Digital Pattern Recognition by Moments," J. ACM, 9 (2), April 1962, 240-258.
9. Bledsoe, W. W., Browning, I., "Pattern Recognition and Reading by Machine, Proc. Eastern Joint Computer Conference, Boston, 1959, 225-232.
10. Lombroso, Cesare, Criminal Man.

REFERENCES CONTINUED

11. Bertillion, Alphonse, The Bertillion System of Identification or
Signaletic Instruction, The Werner Co., Chicago, New York, 1896.

121
PRI 18

MEASUREMENTS BY COMPUTER
OF THE DISTANCES ON AND
ABOUT THE EYES

C. L. BISSON

April 1965

PRI:18

FACIAL RECOGNITION PROJECT

PRELIMINARY INVESTIGATION ON

MEASUREMENTS BY COMPUTER OF THE
DISTANCES ON AND ABOUT THE EYES

C. L. BISSON

1965

Panoramic Research, Inc., 2670 Hanover Street, Palo Alto, California.

MEASUREMENTS

The measurements to be made are the distances between the outside corners of the eyes, and the inside corners of the eyes as indicated by D_o and D_i respectively in Figure 8. Thus the problem for the machine is to locate the coordinates of these points. Some additional points were located in order to find the corners, namely, the top, bottom, left and right sides of the pupil and points along the eyelids as shown in their order of computation in Figure 8. The measurements were made by a CDC 1604A Computer using photographs that had been scanned, digitized and written on magnetic tape.

ASSUMPTIONS

The method made two assumptions. The first was knowledge of the general location of the horizontal section across the eyes which must include vertically about three pupil diameters, one above and one below the pupil. The second assumption made was that a rough estimate of the pupil diameter was known.

RESULTS

Of the seven sections that were scanned Figures 1 to 7, only four were measured and the located coordinates are shown in Figure 1b to 4b. It should be noted that the program was unable to locate the right eyes in Figures 3b and 4b. The right side of Figure 4b was rejected by the program and in Figure 3b inadvertently not enough data above the eye had been scanned.

The method does not work on the outer corners but the locations of the inside corners have been computed in each case that the pupil had been located. (cf Figures 1b to 4b.) The inability to locate the outer corners is due to the lack of definite black-white transitions caused by the direction of incident light, graininess in the photograph, etc. In poor lighting the

outside corners are not well defined and their precise location by humans would be difficult and more difficult for a machine.

EQUIPMENT USED

The flow of data handling from source photograph to the final form of the data (magnetic tape) is shown in Figure 9. The output from the TELauto-graph scanner had to be modified to be compatible with the A/D converter and the necessary circuitry designed and installed by Mr. B. T. Harrell of PRI is shown in Figure 10. The scanner has a vertical density of 100 lines/inch and the sampling rate of the A/D converter was chosen so that the horizontal sampling density was also 100 samples/inch. The effective spot size of the scanner with the above modification was about 2-3 sample points. The A/D converter output was 6 bits/sample point (i. e. 64 gray-levels) and the final data format was about 800, 6 bit characters per scan line and was written as one record/scan line.

METHOD OF MEASURING

The method of measuring by the program was divided into four steps.

1. Location of the pupil coordinates.
2. Location at the points along the upper and lower eye lids as shown in Figure 8.
3. Fitting parabolas to the points along the upper and lower portions of the eye lid.
4. Computing the intersection of the upper and lower parabolas.

Steps 1 and 2 were achieved using the same technique of searching in a vertical (horizontal) direction for the maxima of the differences in average intensity above and below (left and right) each point in a vertical (horizontal) strip.

As an example, suppose that one wished to locate the bottom of the black area shown in Figure 11. Then by fixing j one computes for each i the difference

$$R_i = \bar{I}_A - \bar{I}_B \quad (\text{White} = 0 \leq I \leq 63 = \text{Black})$$

where $\bar{I}_A$ is the average intensity over an area above i and $\bar{I}_B$ is the average intensity over an area below i . The boundary at $i = k$ indicated in Figure 11 is located by starting to search the values of R_i in the area known to be white and searching the direction toward the black for the first proper maximum (R_k) occurring at $i = k$. This technique works best when one can choose the size of the areas to fix approximately the areas defined by the video data.

The curve fitting was done by a simple least-squares polynomial fit and in the case the mean-square error was too large, smoothing was attempted by the deletion at data points only the no gap filling was done. The deletion of points was done by computing the mean μ and standard deviation σ of the error and deleting those points whose errors were greater than $\mu + \sigma$. The data points that were found, their fitted curves and the deleted data points are shown in Figures 12 to 19.

SUMMARY

The methods used have located the inside corners in 6 out of 8 cases but failed when trying to locate the outside corners and it is felt that this approach would not suffice. The search for the pupil made an unnecessary rejection (cf right side of Figure 4b) but this could be remedied. Other shortcomings are; long shadows over the eyes and the narrowing of the eye lids. The latter could be overcome but making measurements among deep shadows will require more sophisticated techniques. As to the accuracy of the measurements the reader is asked to compare the source photographs with the accompanying reproductions with the computed coordinates shown. (Figure 1 to 4)

If accurate measurements are to be made in larger quantities effort must be given reducing the computing time, reducing the effective spot size of the scanner and perhaps increasing the sampling density to say, 200 samples/inch. In making measurements one is not trying to throw information away but to use as many data points as possible for each measurement.



FIGURE 1a

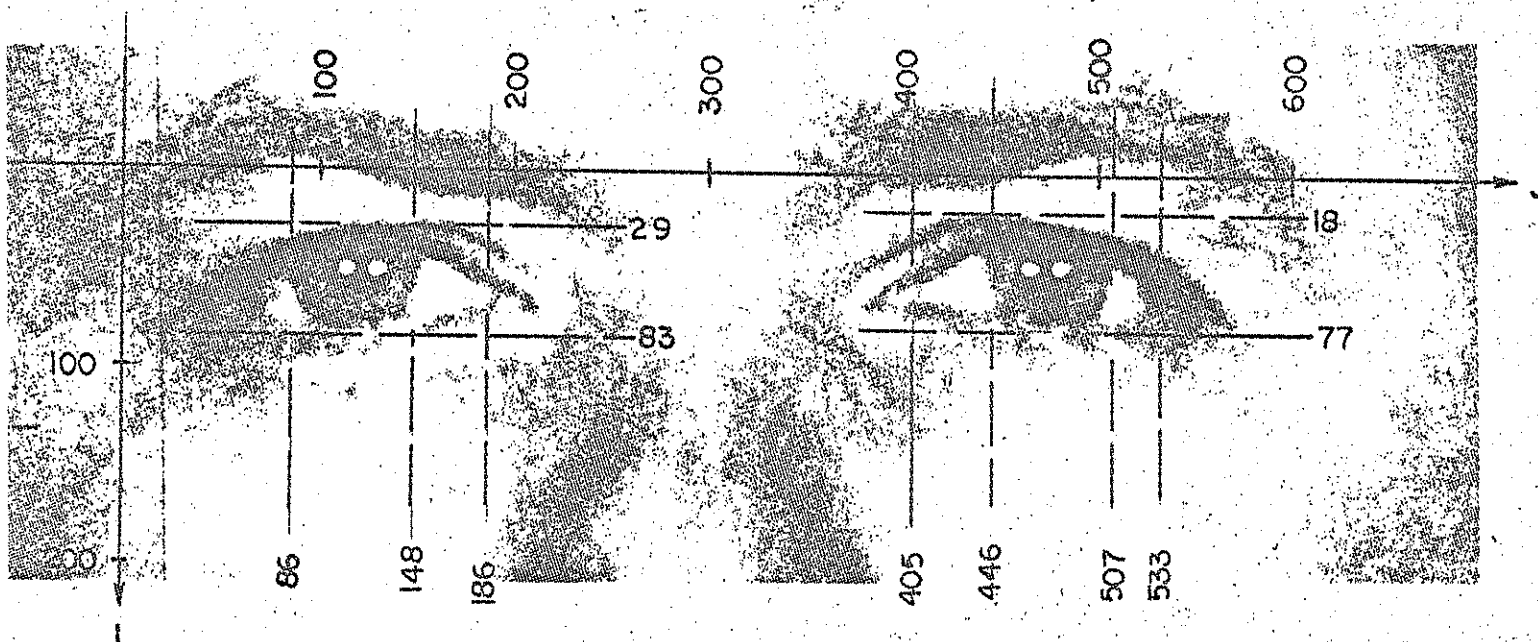


FIGURE 1b



FIGURE 2a

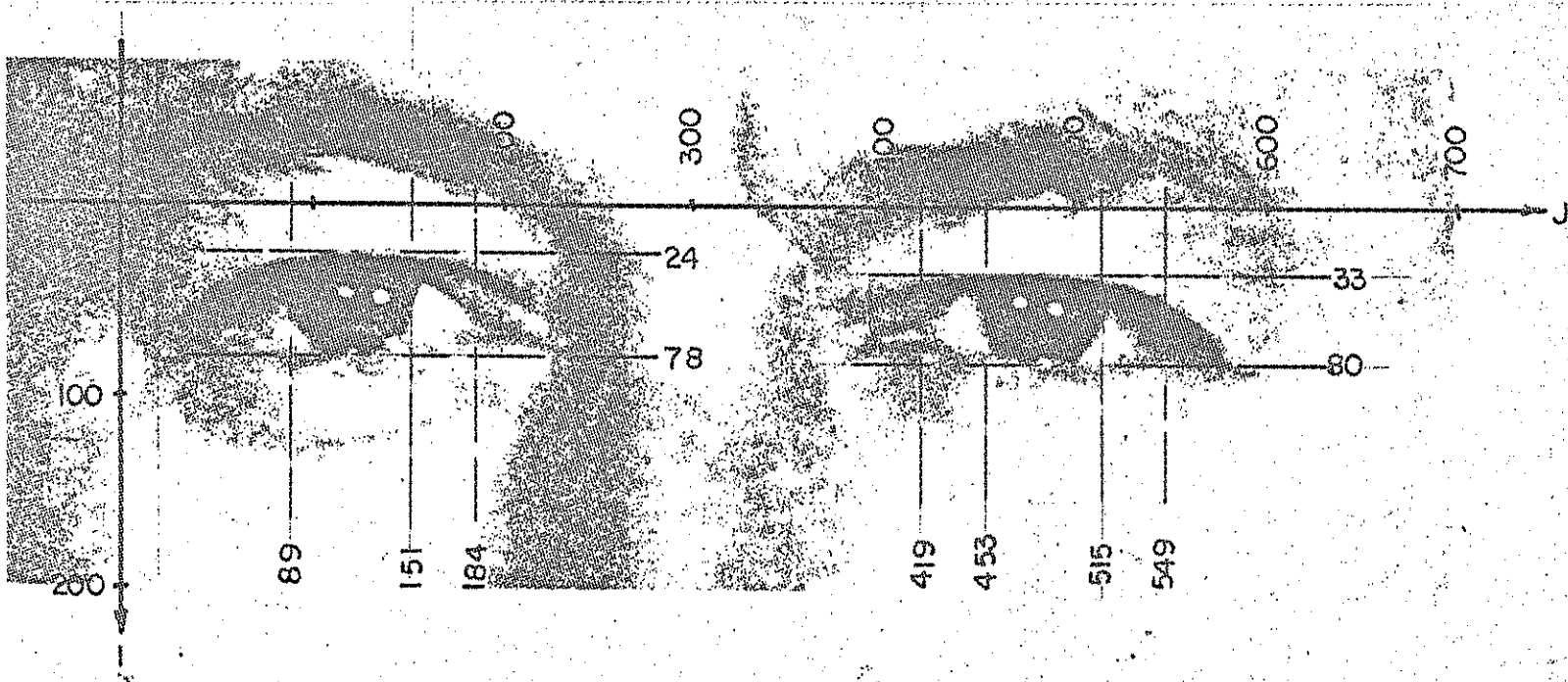


FIGURE 2b



FIGURE 3a

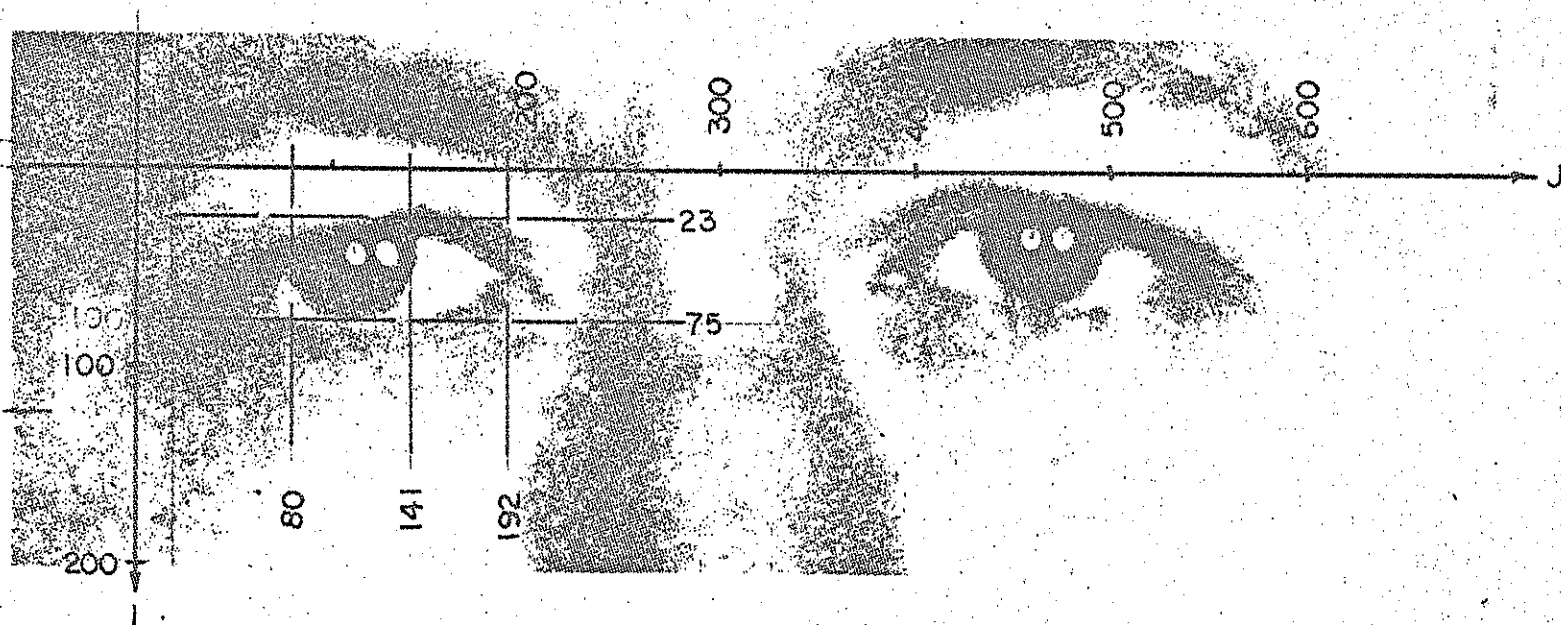


FIGURE 3b



FIGURE 4a

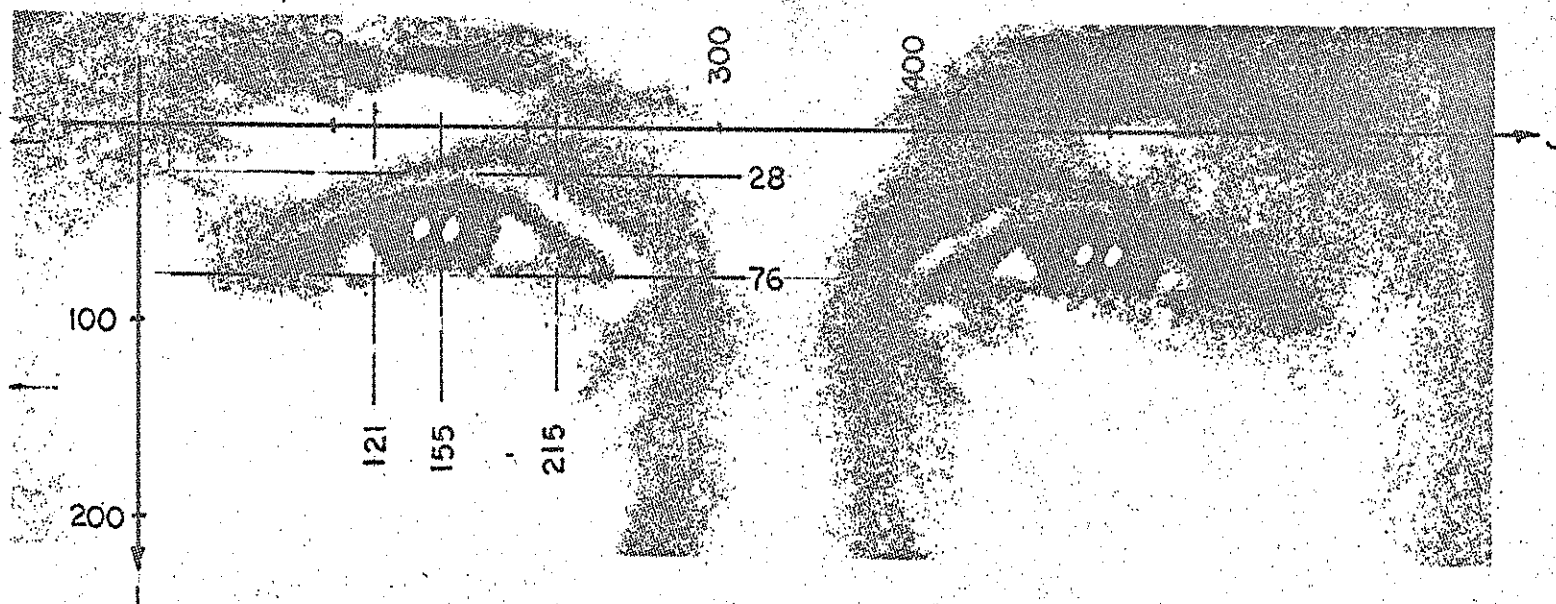


FIGURE 4b



FIGURE 5



FIGURE 6



FIGURE 7

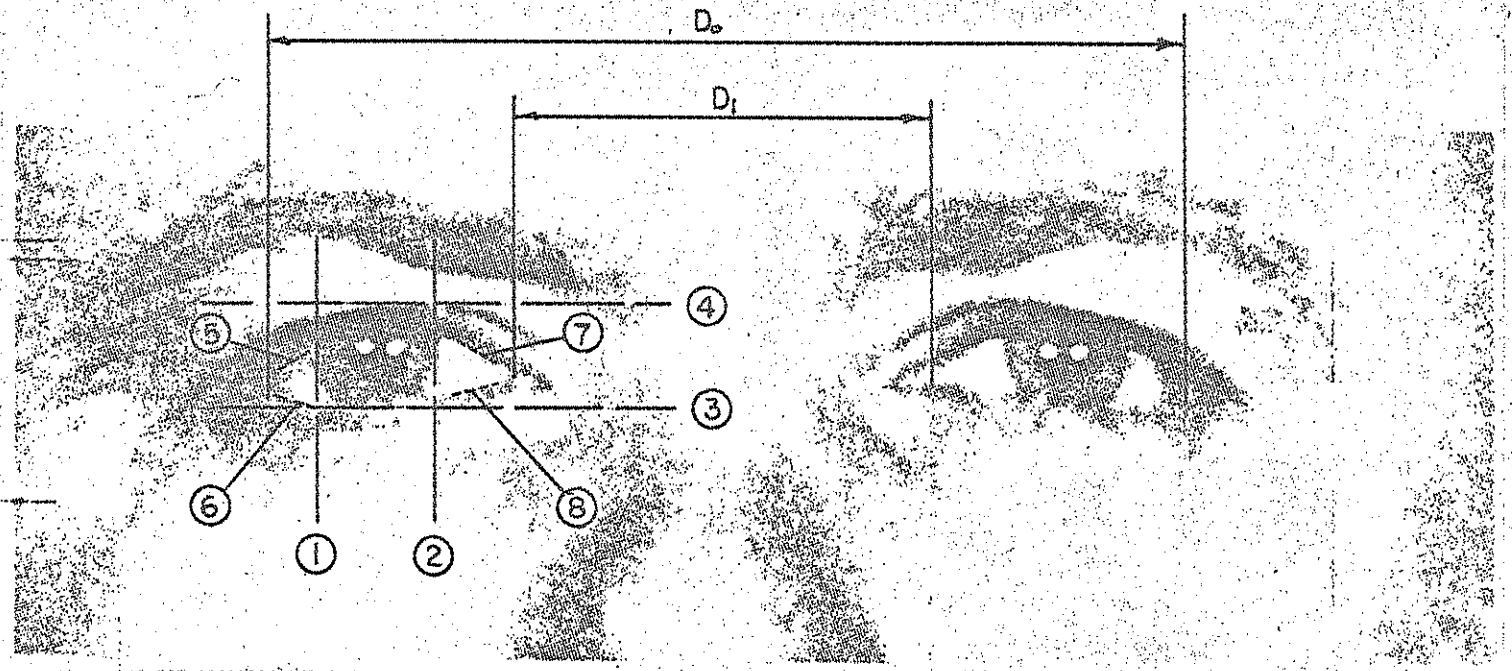
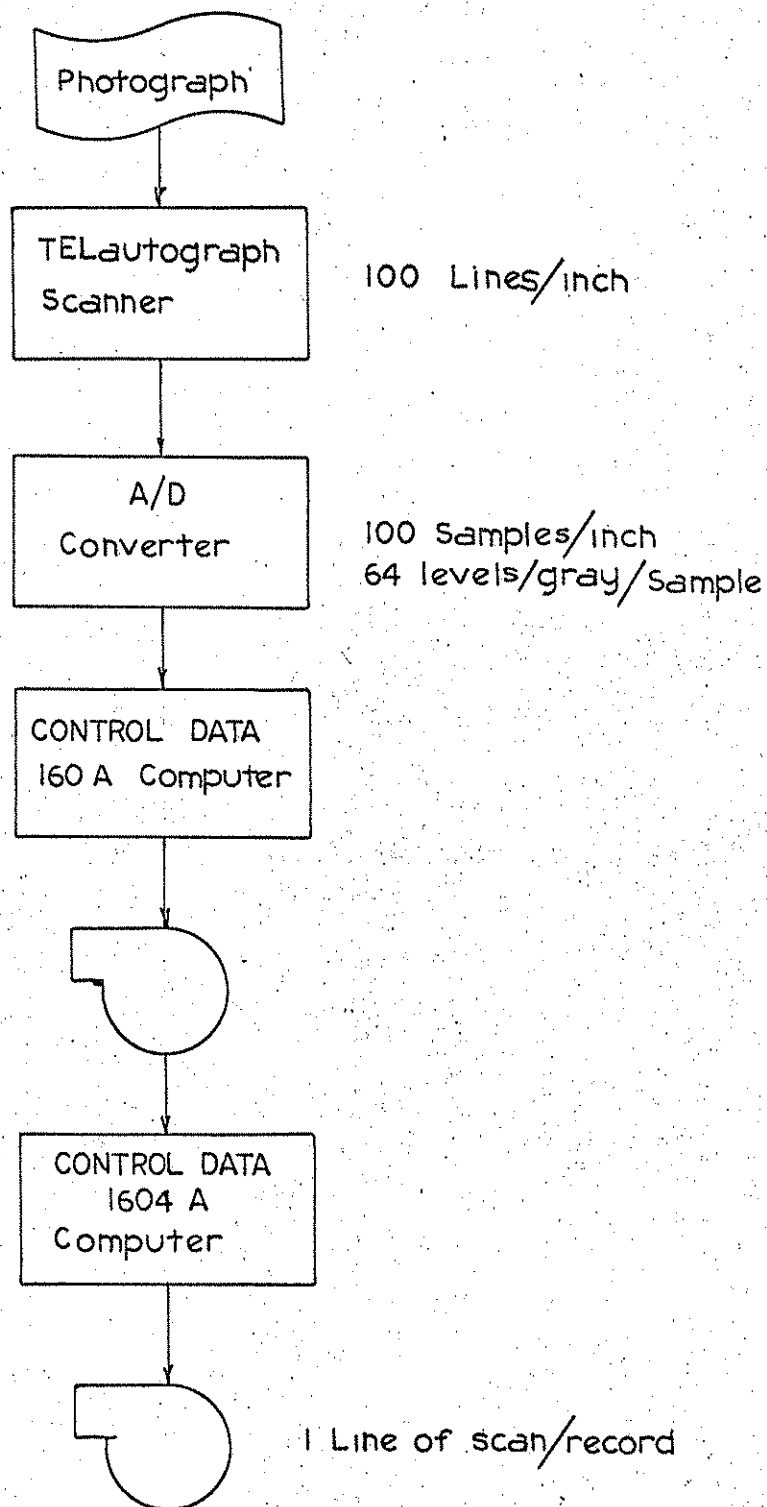


FIGURE 8



FLOW OF DATA

FIGURE 9

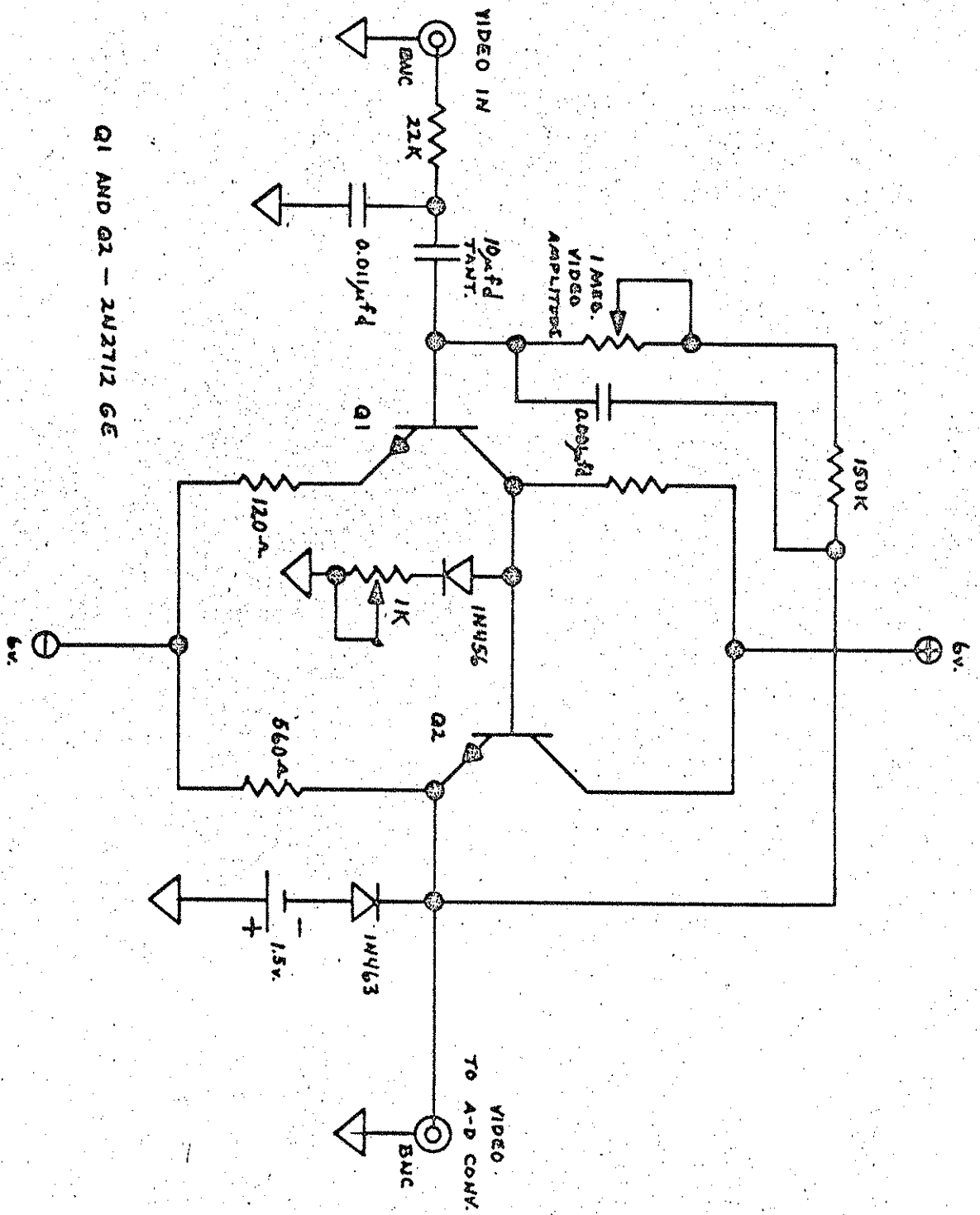
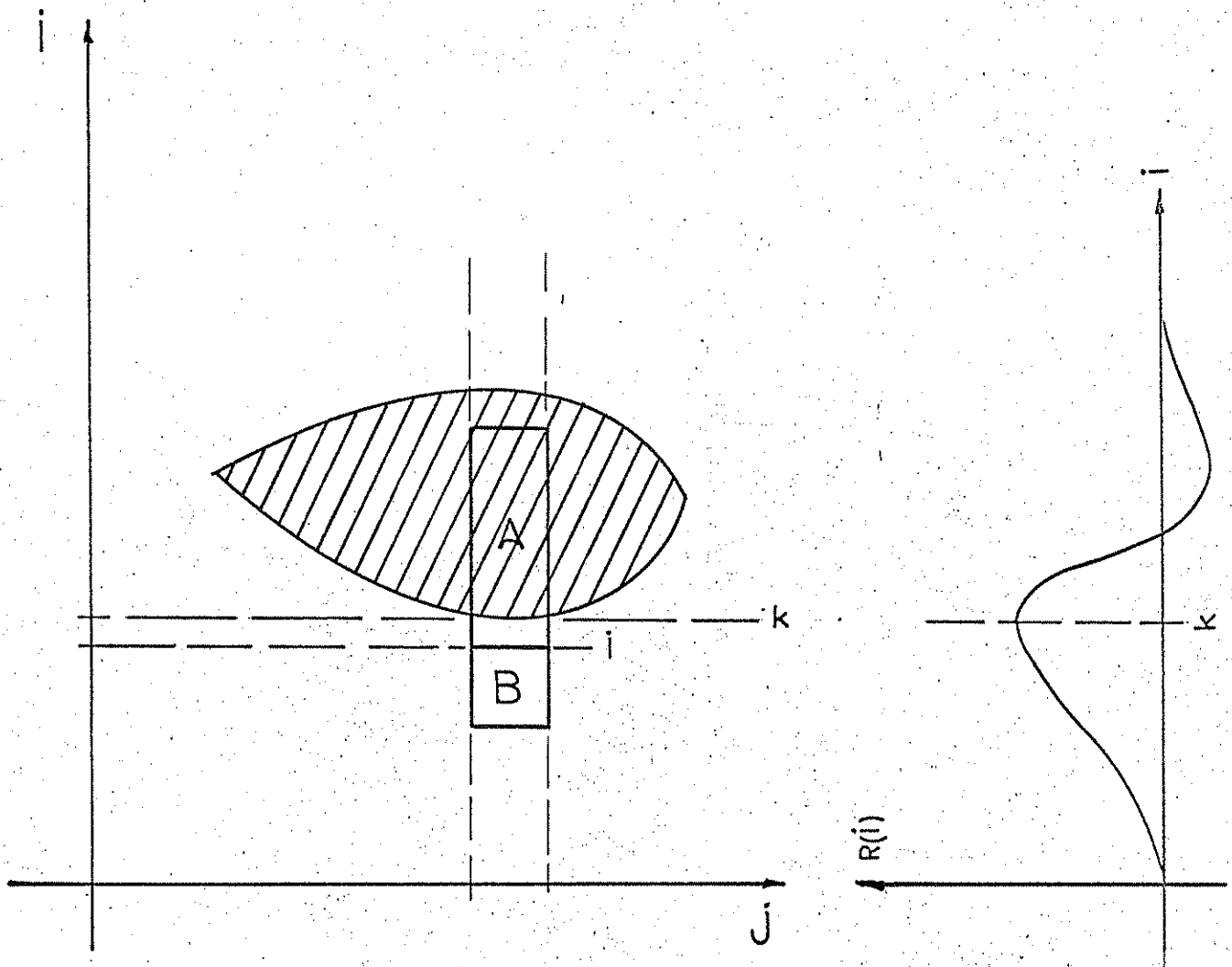


FIGURE 10

INTERFACE BETWEEN CDC A-D CONV. AND TELETGRAPHY SCANNER

13.7.24.

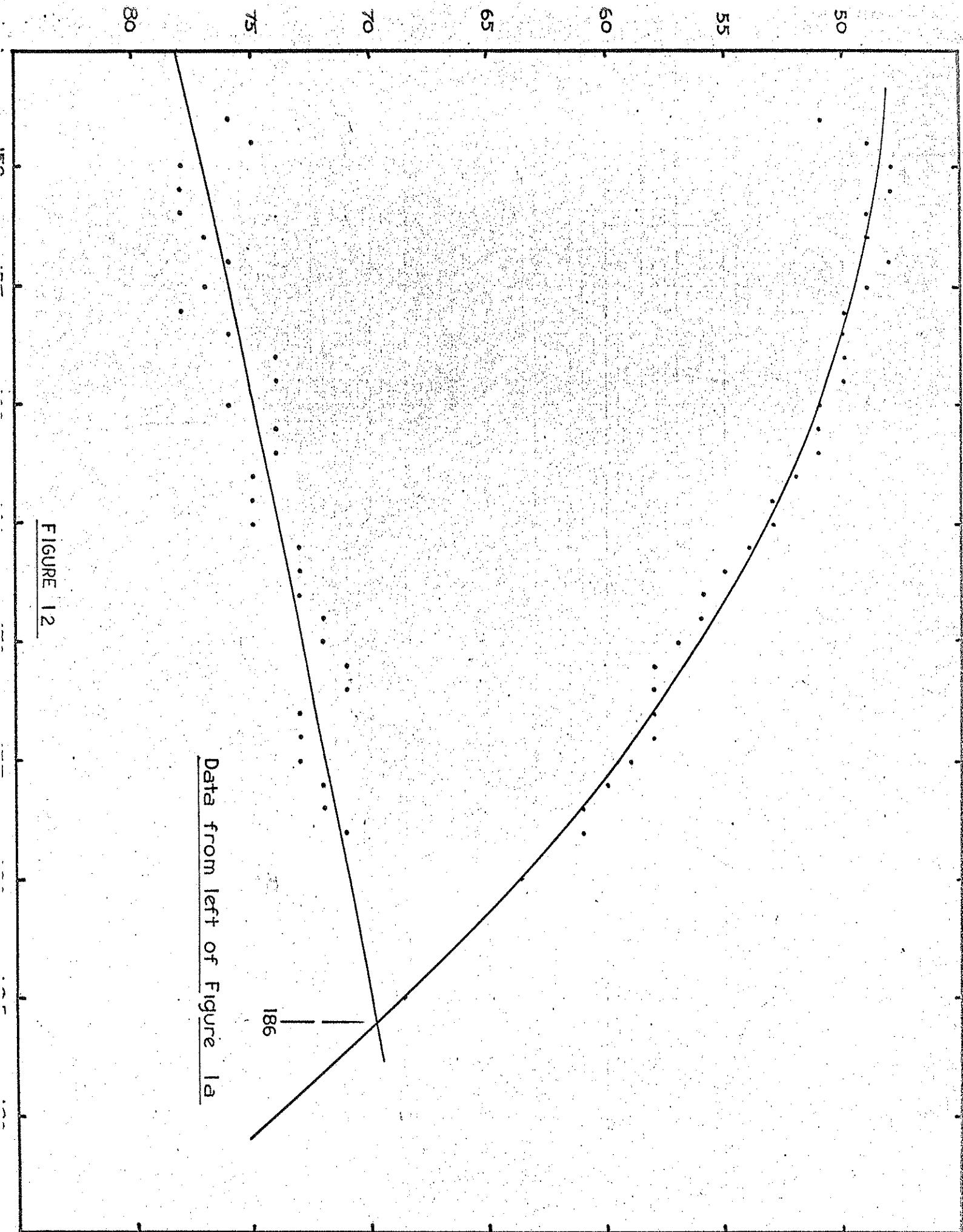


$$R_i = \bar{I}_A - \bar{I}_B \quad \text{where}$$

$\bar{I}_A$ = Average intensity over area A

$\bar{I}_B$ = Average intensity over area B

FIGURE II



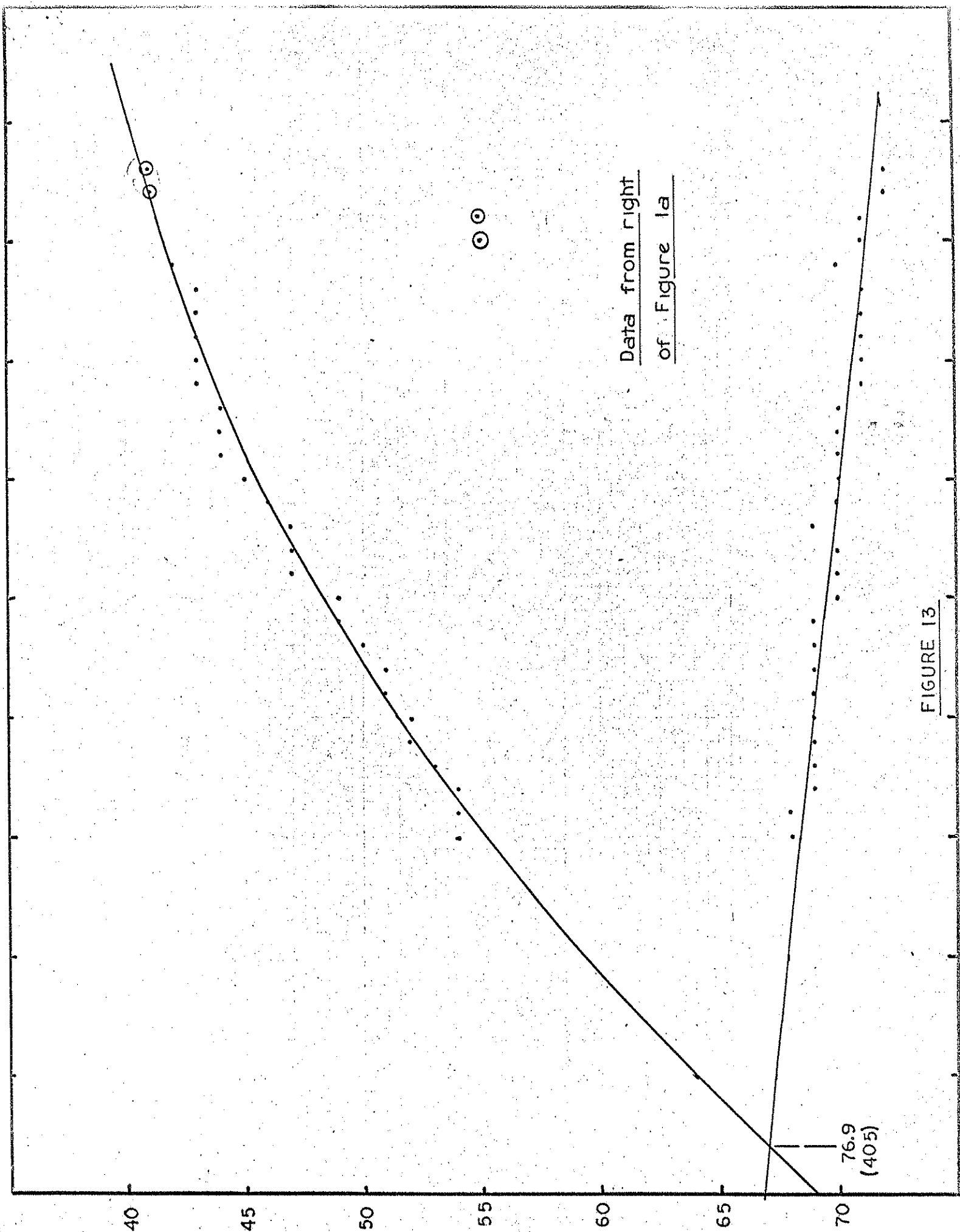
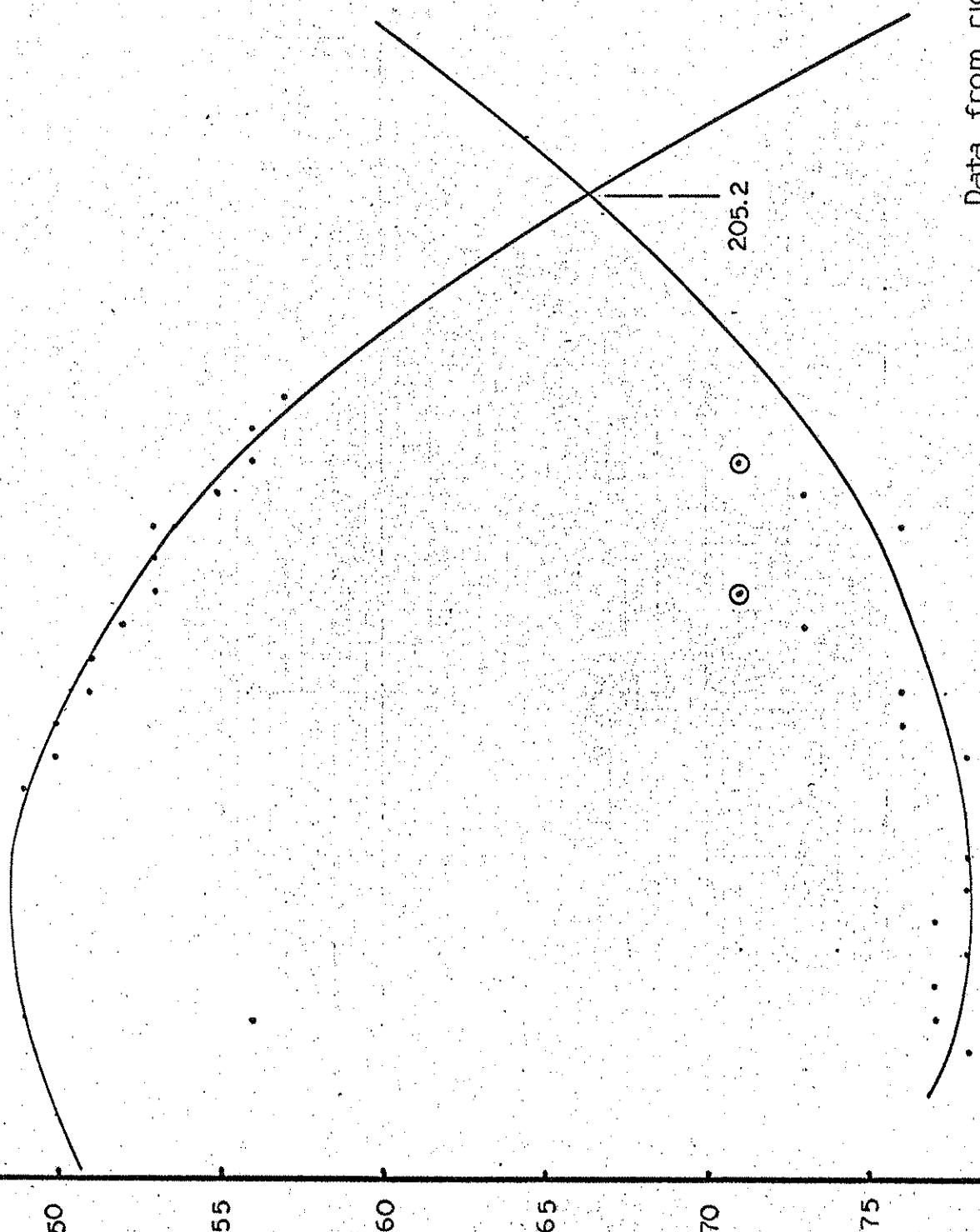
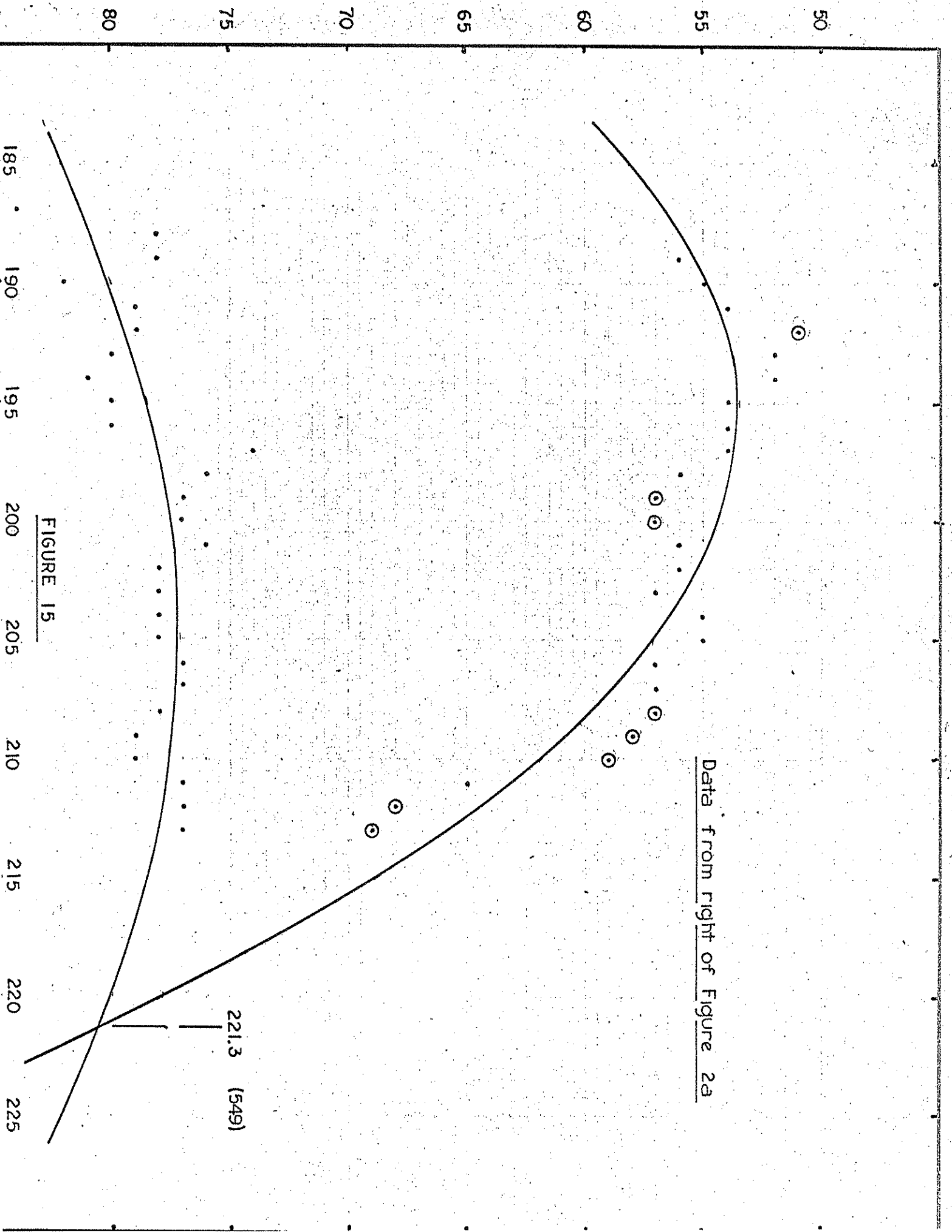


FIGURE 13



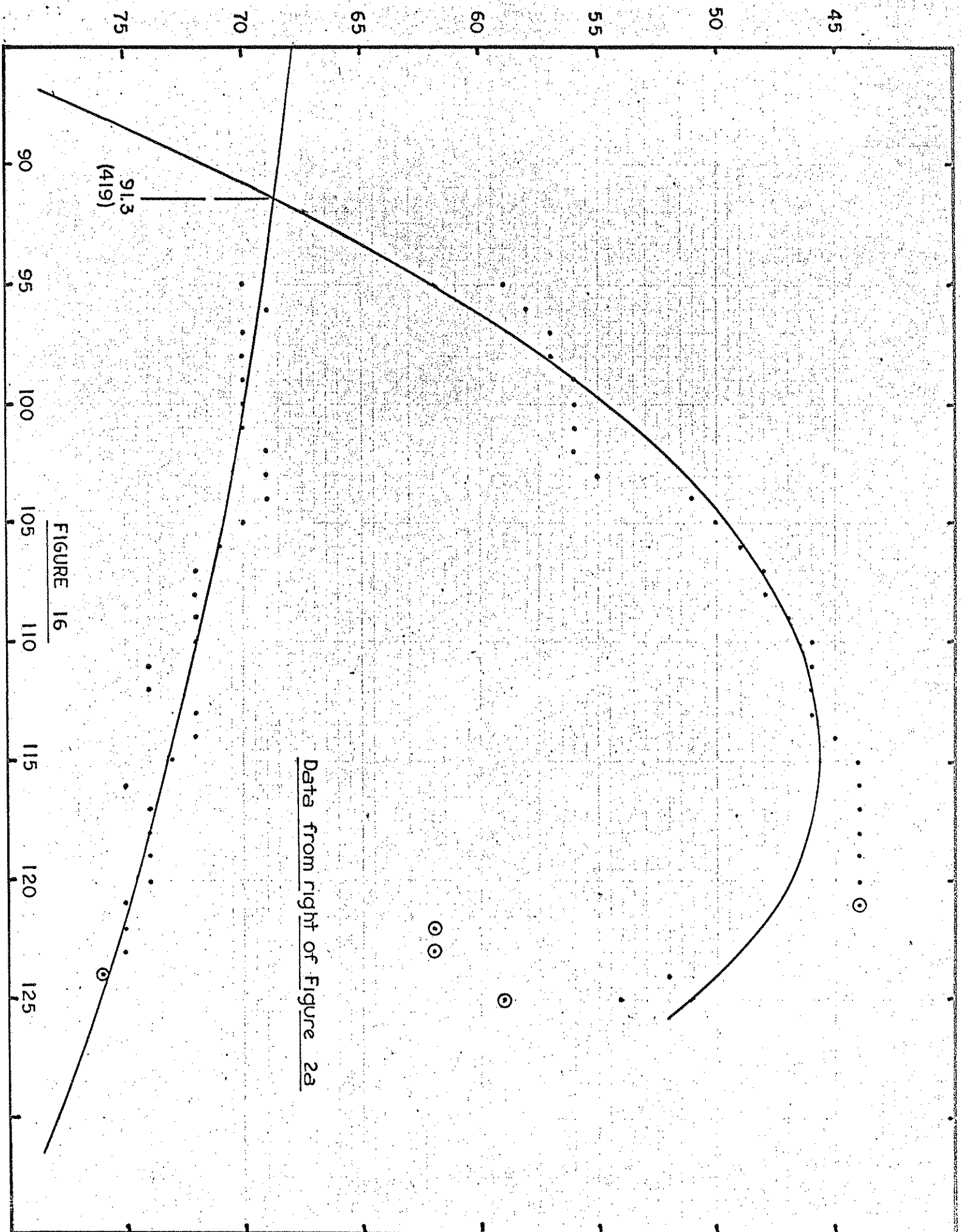
Data from right of Figure 1a

FIGURE 14



Data from right of Figure 2a

FIGURE 15

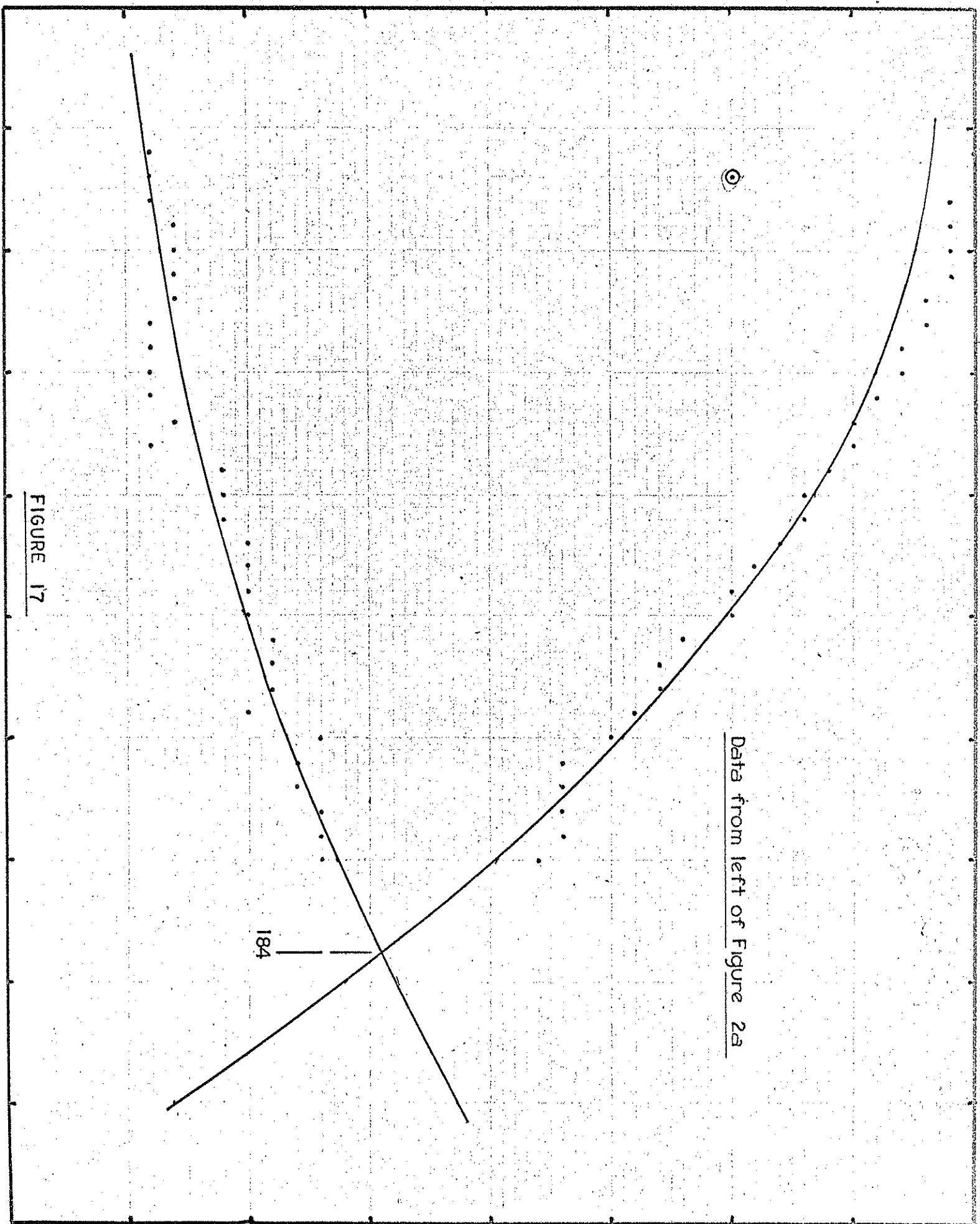


45
50
55
60
65
70
75

Data from left of Figure 2a

184

FIGURE 17



Data from left of Figure 3a

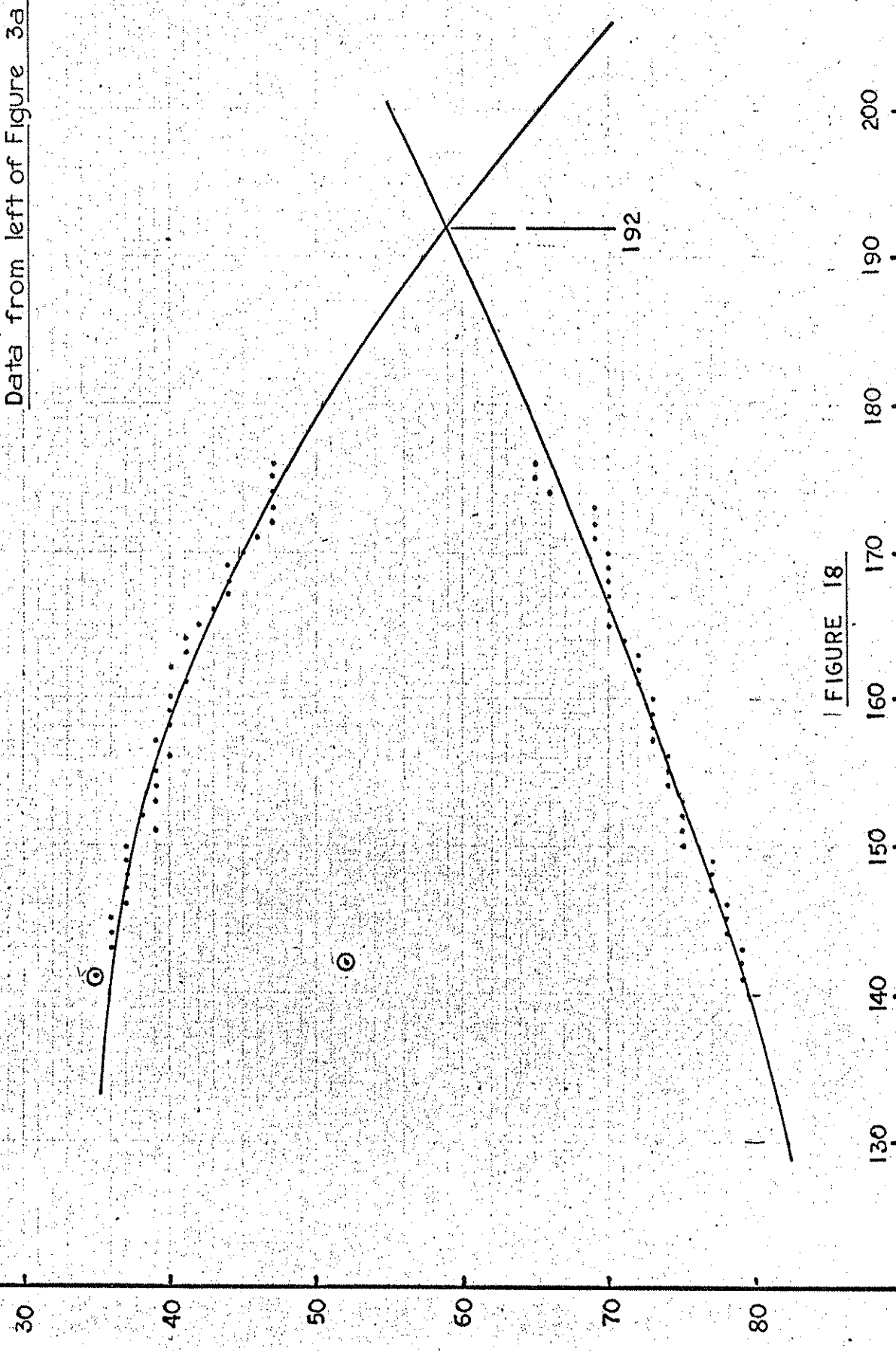
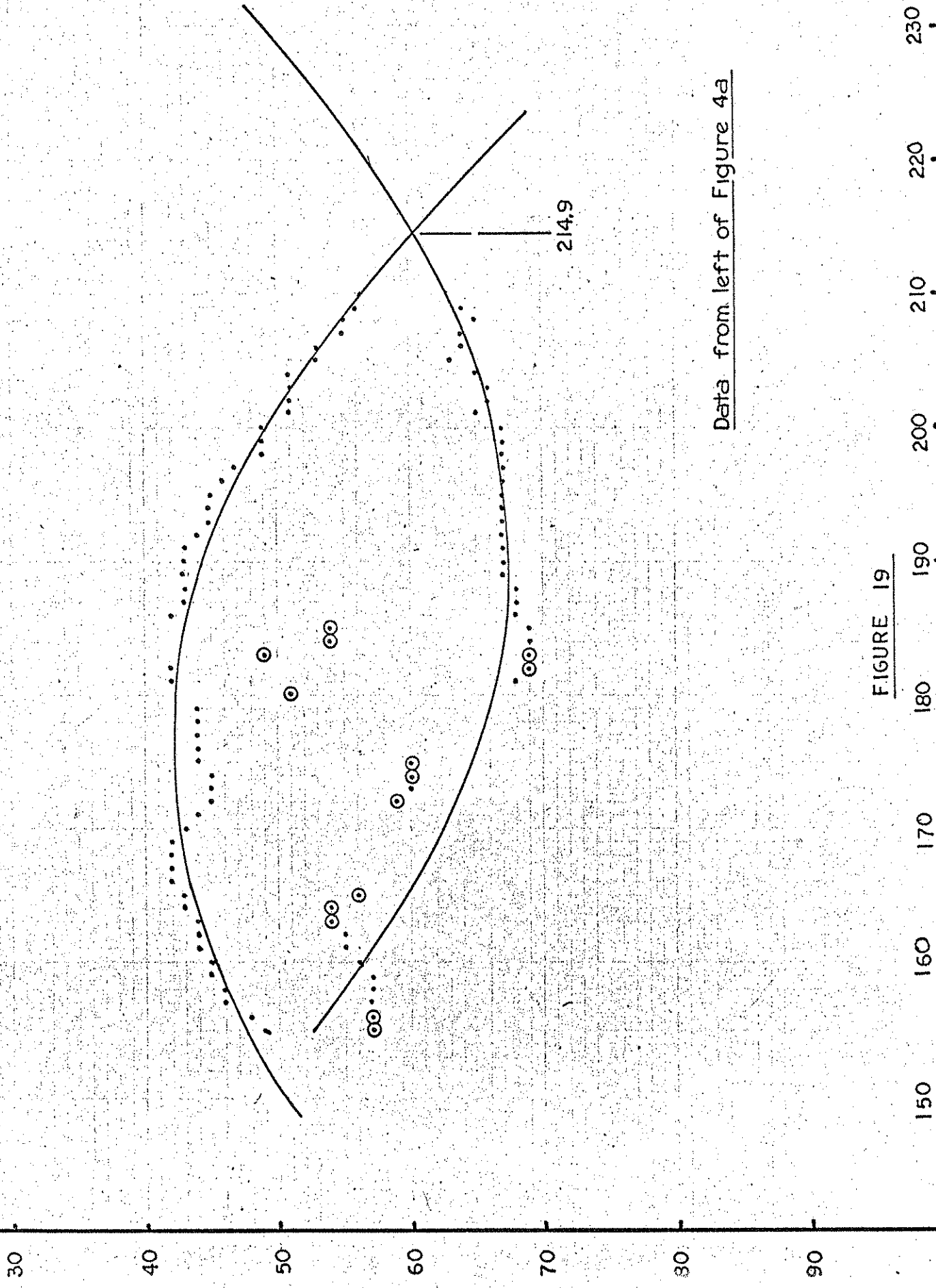


FIGURE 18



Data from left of Figure 4a

FIGURE 19

Phone *PRI 20*
June 1965
File Box

LOCATION OF SOME FACIAL FEATURES BY COMPUTER

C. L. Bisson

Automatic measurement by a computer of distances on facial photographs might well proceed in two distinct steps.

1. Finding the general location of the features to be measured.
2. Precise measurements of features by automatic or semi-automatic means.

This report concerns some work that has been done to achieve step 1 and the status of present efforts.

The general features to be located are the sides and top of the head, location of the forehead, eyes, tip of nose, line along mouth, and bottom of the chin. These features are located by horizontal and vertical lines as indicated in Figure 1 along with their order of computation.

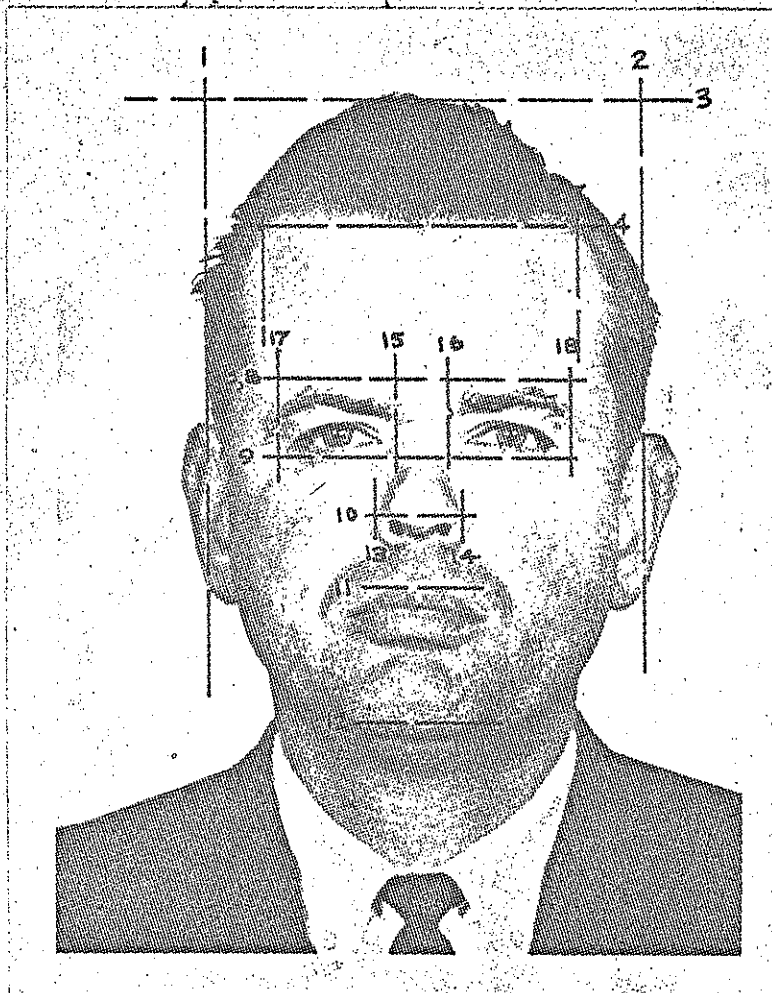


Figure 1.

Features to be located
and order of computation.

Method

The method used to locate a boundary defined by a black-white (or white-black) transition is merely searching for the maxima of the differences in average intensity over certain areas. As an example suppose that one wished to locate the bottom of the black area shown in Figure 2. Then by fixing j one computes for each i the difference

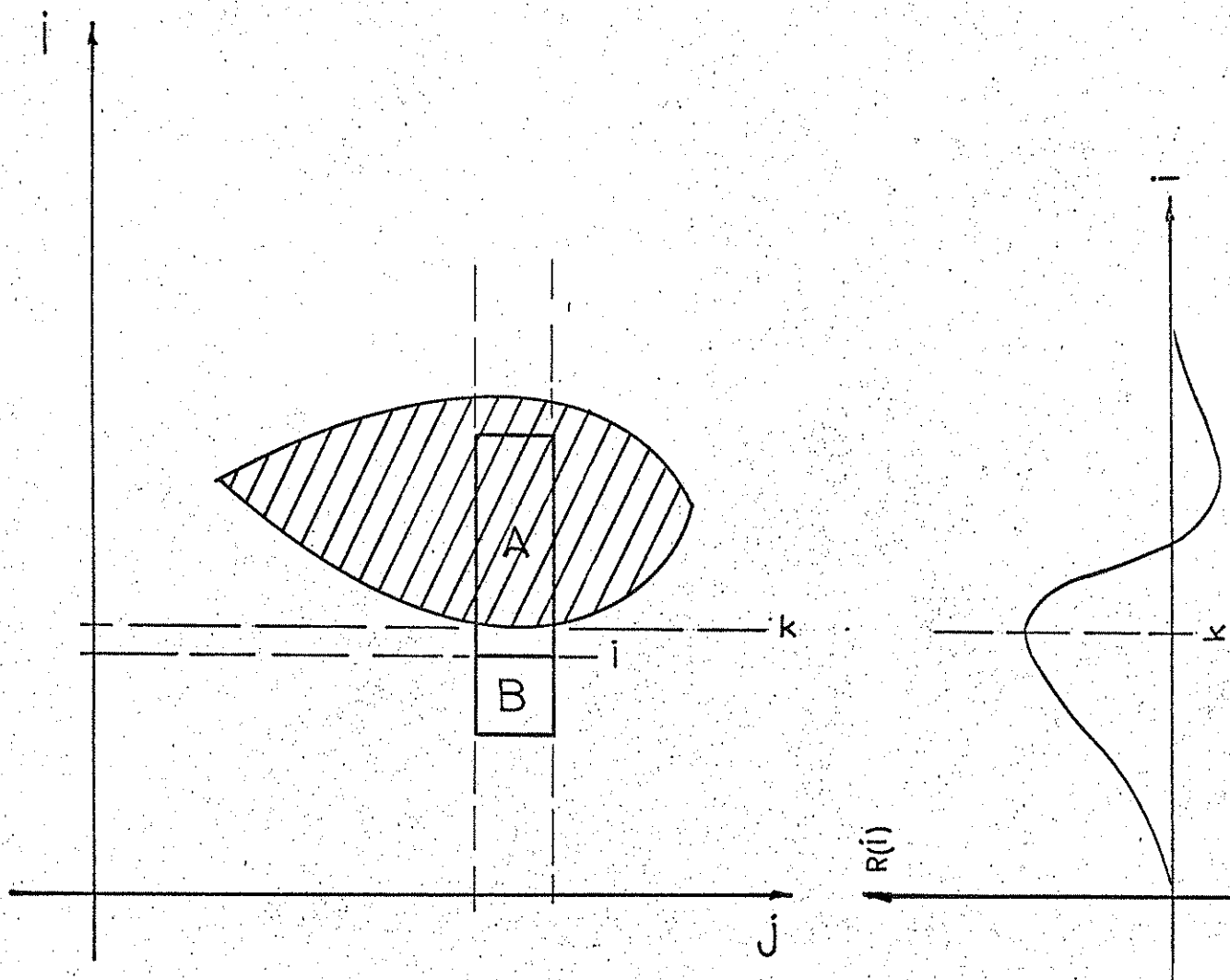
$$R_i = T_A - T_B \quad (\text{here white} = 0 \leq i \leq 15 = \text{black})$$

where T_A is the average intensity over an area above i and T_B is the average intensity over an area below i . The boundary at $i = k$ indicated in Figure 2 is located by starting to search the values of R_i in the area known to be white and searching in the direction toward the black for the first proper maximum (R_k) occurring at $i = k$. This technique works best when one can choose the size of the areas to fit approximately the areas defined by the video data. This choice of area size is sometimes possible by using ratios of facial distances such as the width of eyes to the width of head, etc.

Data

The treatment of data has gone through several sizes of photographs and size reduction by computer so as to be in a convenient form with regard to access speed and fidelity. The final form of the present data is such that a 4" x 5" photograph is well represented by 18,500 -- 48 bit words of storage. This form of the data is achieved by the following steps.

1. The source photograph is scanned by a modified Telautograph scanner with the horizontal and vertical sample densities at 100 samples/inch.
2. An analog-to-digital conversion in which each analog video sample is converted to an 8-bit quantity (i. e. 256 levels of gray) and written on magnetic tape.
3. A preliminary pass on a CDC 1604A computer in which each sample is reduced to only 4 bits (16 levels/gray) and packed 12 samples/word and written as 500, 33 word records on magnetic tape.



$$R_i = \bar{I}_A - \bar{I}_B \quad \text{where}$$

$\bar{I}_A$ = Average intensity over area A

$\bar{I}_B$ = Average intensity over area B

FIGURE 2

A sample printout of data from the area around the left eye of the photograph in Figure 6 is shown in Figure 7. Sixteen gray levels are represented by White = 0, 1, ..., 9, A, B, C, D, E, F = Black.

Results to Date

Several photographs have been scanned for processing by this method (See Figures 3-6). In particular one photograph has been processed and the features (as defined in Figure 1) have been located by the computer. These results are shown in Figure 6. The horizontal and vertical lines in Figure 6 indicate the x and y coordinates obtained by the computer to delineate the various features.

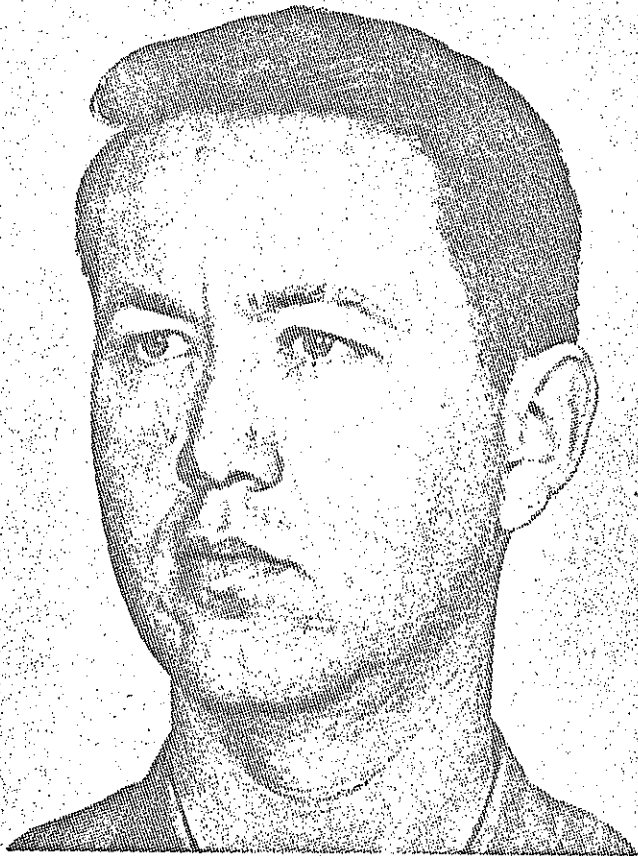


FIGURE 3



FIGURE 4



FIGURE 5

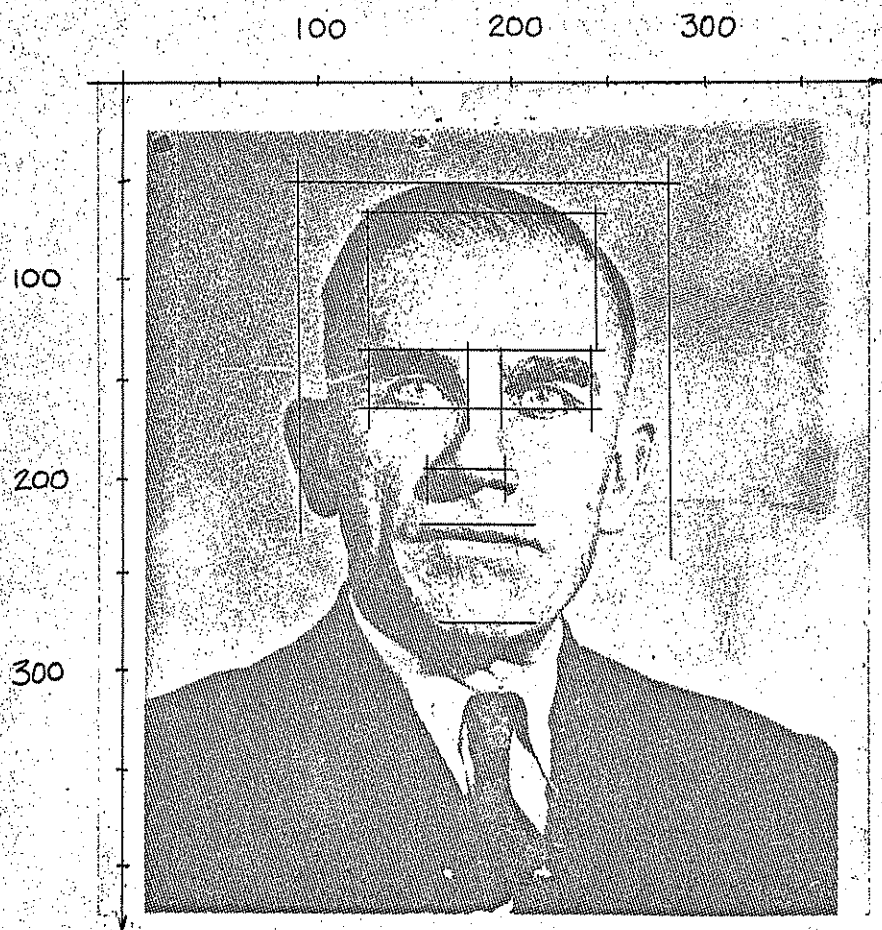


FIGURE 6
FEATURES LOCATED BY MACHINE

David B. B.

Formal Education and Training

PRI: 20

1965

LOCATION OF SOME FACIAL FEATURES BY COMPUTER

C. L. Bisson

Automatic measurement by a computer of distances on facial photographs might well proceed in two distinct steps.

1. Finding the general location of the features to be measured.
2. Precise measurements of features by automatic or semi-automatic means.

This report concerns some work that has been done to achieve step 1 and the status of present efforts.

The general features to be located are the sides and top of the head, location of the forehead, eyes, tip of nose, line along mouth, and bottom of the chin. These features are located by horizontal and vertical lines as indicated in Figure 1 along with their order of computation.

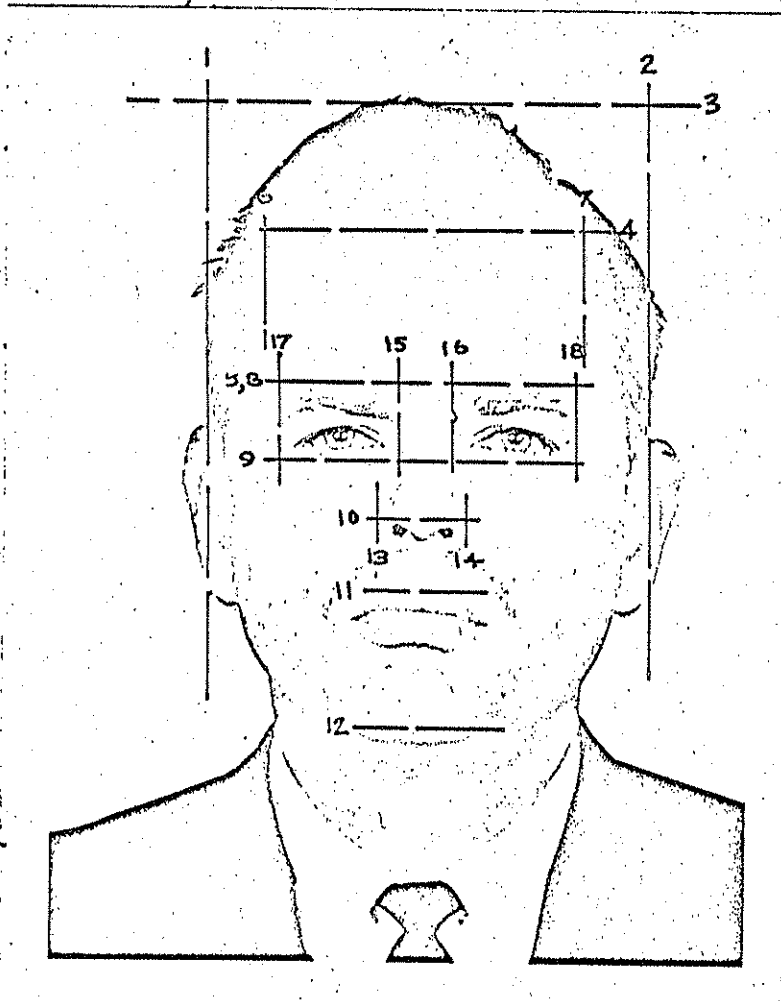


Figure 1.

Features to be located
and order of computation.

Method

The method used to locate a boundary defined by a black-white (or white-black) transition is merely searching for the maxima of the differences in average intensity over certain areas. As an example suppose that one wished to locate the bottom of the black area shown in Figure 2. Then by fixing j one computes for each i the difference

$$R_i = \bar{I}_A - \bar{I}_B \quad (\text{here white} = 0 \leq I \leq 15 = \text{black})$$

where $\bar{I}_A$ is the average intensity over an area above i and $\bar{I}_B$ is the average intensity over an area below i . The boundary at $i = k$ indicated in Figure 2 is located by starting to search the values of R_i in the area known to be white and searching in the direction toward the black for the first proper maximum (R_k) occurring at $i = k$. This technique works best when one can choose the size of the areas to fit approximately the areas defined by the video data. This choice of area size is sometimes possible by using ratios of facial distances such as the width of eyes to the width of head, etc.

Data

The treatment of data has gone through several sizes of photographs and size reduction by computer so as to be in a convenient form with regard to access speed and fidelity. The final form of the present data is such that a 4" x 5" photograph is well represented by 18,500 -- 48 bit words of storage. This form of the data is achieved by the following steps.

1. The source photograph is scanned by a modified Telautograph scanner with the horizontal and vertical sample densities at 100 samples/inch.
2. An analog-to-digital conversion in which each analog video sample is converted to an 8-bit quantity (i. e. 256 levels of gray) and written on magnetic tape.
3. A preliminary pass on a CDC 1604A computer in which each sample is reduced to only 4 bits (16 levels/gray) and packed 12 samples/word and written as 500, 33 word records on magnetic tape.

Method

The method used to locate a boundary defined by a black-white (or white-black) transition is merely searching for the maxima of the differences in average intensity over certain areas. As an example suppose that one wished to locate the bottom of the black area shown in Figure 2. Then by fixing j one computes for each i the difference

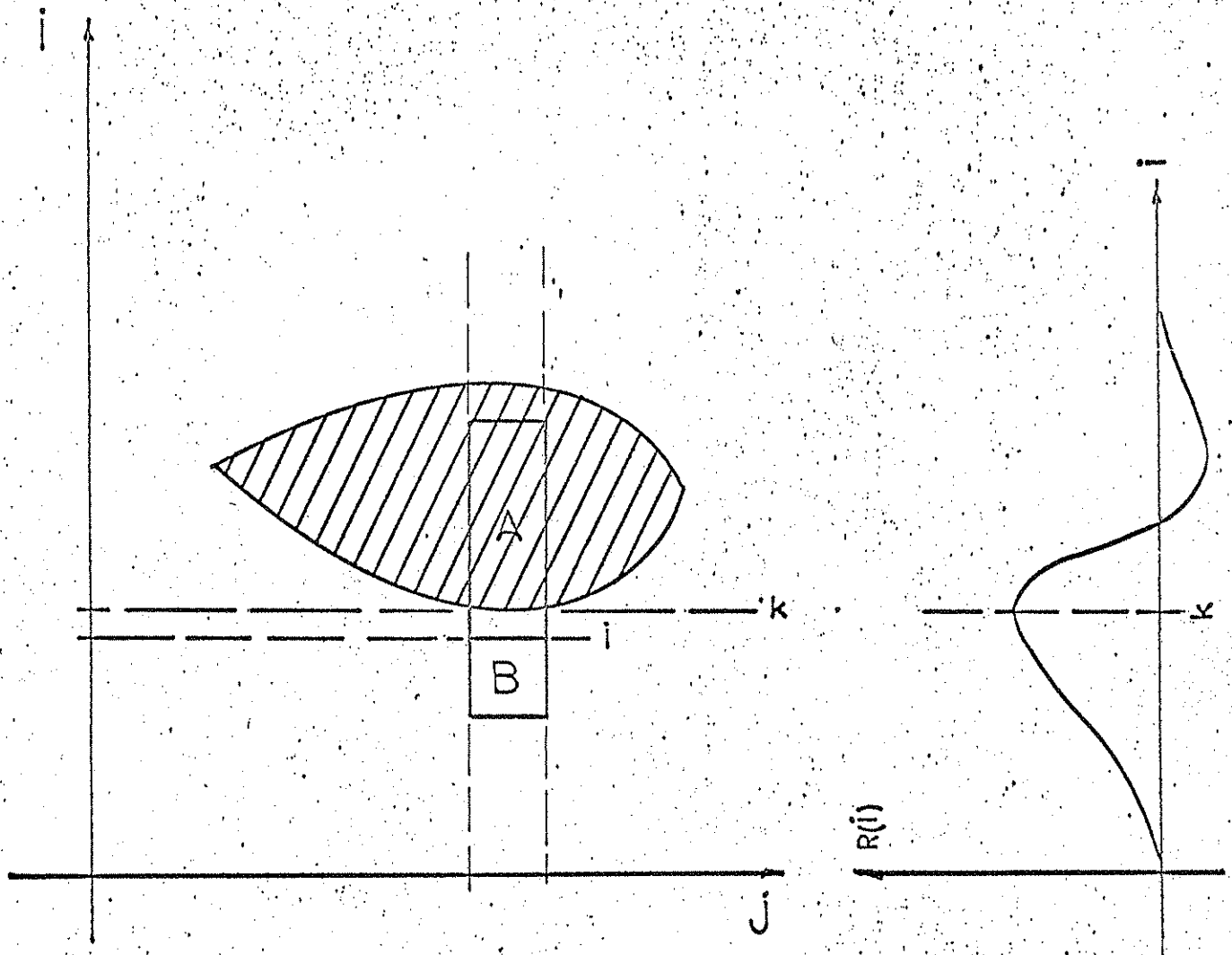
$$R_i = \bar{I}_A - \bar{I}_B \quad (\text{here white} = 0 \leq I \leq 15 = \text{black})$$

where $\bar{I}_A$ is the average intensity over an area above i and $\bar{I}_B$ is the average intensity over an area below i . The boundary at $i = k$ indicated in Figure 2 is located by starting to search the values of R_i in the area known to be white and searching in the direction toward the black for the first proper maximum (R_k) occurring at $i = k$. This technique works best when one can choose the size of the areas to fit approximately the areas defined by the video data. This choice of area size is sometimes possible by using ratios of facial distances such as the width of eyes to the width of head, etc.

Data

The treatment of data has gone through several sizes of photographs and size reduction by computer so as to be in a convenient form with regard to access speed and fidelity. The final form of the present data is such that a 4" x 5" photograph is well represented by 18,500 ~ 48 bit words of storage. This form of the data is achieved by the following steps.

1. The source photograph is scanned by a modified Telautograph scanner with the horizontal and vertical sample densities at 100 samples/inch.
2. An analog-to-digital conversion in which each analog video sample is converted to an 8-bit quantity (i. e. 256 levels of gray) and written on magnetic tape.
3. A preliminary pass on a CDC 1604A computer in which each sample is reduced to only 4 bits (16 levels/gray) and packed 12 samples/word and written as 500, 33 word records on magnetic tape.



$$R_i = \bar{I}_A - \bar{I}_B \quad \text{where}$$

$\bar{I}_A$ = Average intensity over area A

$\bar{I}_B$ = Average intensity over area B

FIGURE 2

A sample printout of data from the area around the left eye of the photograph in Figure 6 is shown in Figure 7. Sixteen gray levels are represented by White = 0, 1, ..., 9, A, B, C, D, E, F = Black.

Results to Date

Several photographs have been scanned for processing by this method (See Figures 3-6). In particular one photograph has been processed and the features (as defined in Figure 1) have been located by the computer. These results are shown in Figure 6. The horizontal and vertical lines in Figure 6 indicate the x and y coordinates obtained by the computer to delineate the various features.

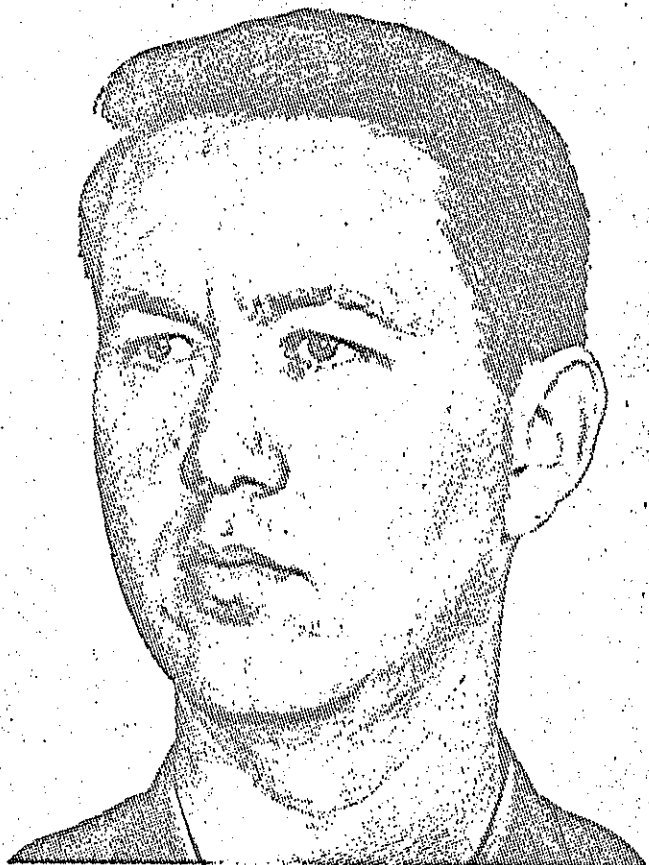


FIGURE 3

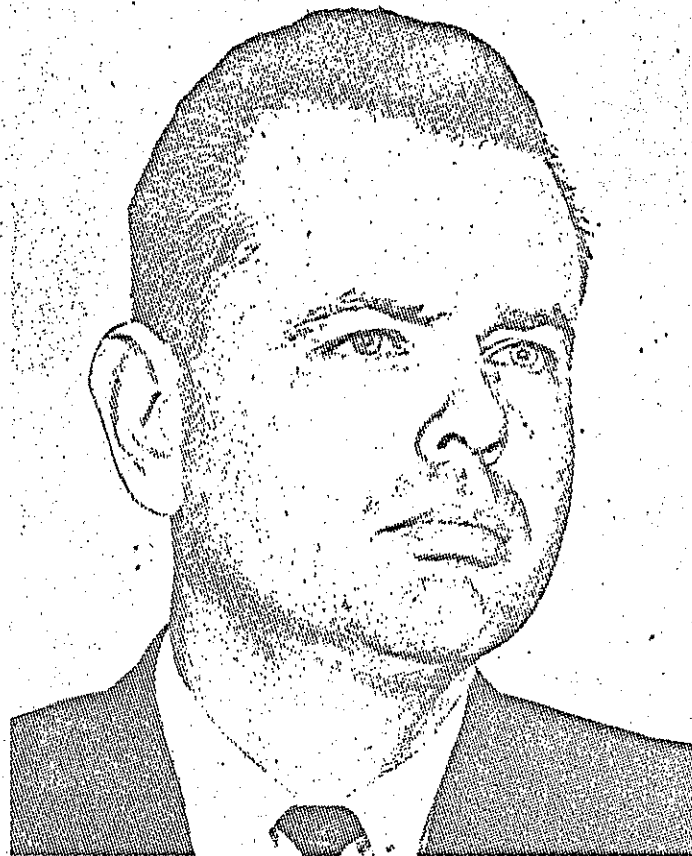


FIGURE 4



FIGURE 5

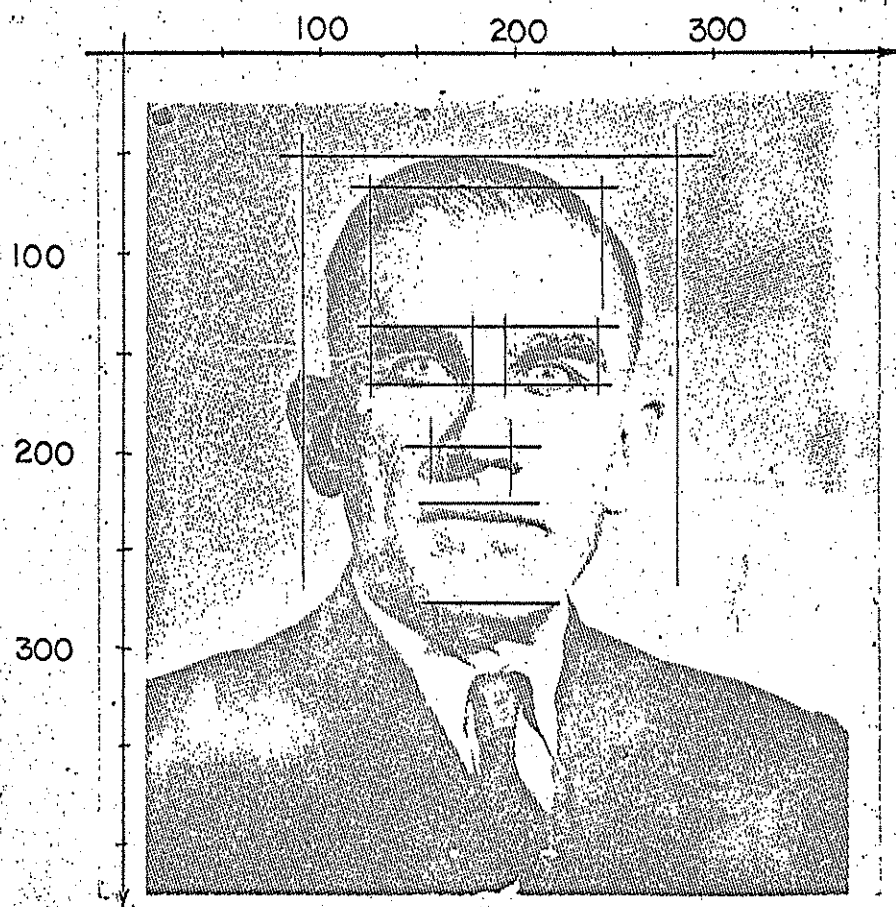


FIGURE 6
FEATURES LOCATED BY MACHINE

[illegible]

INTERNAL MEMORANDUM

19

MR.

Return to Bled soe

RWS

AIR-TARGETS, AIR-BURST PROBLEM

I

SEMINAR TALK

Notes

Oct 16, 1953
AAA Meeting
5100

Targets

1. point targets : Get prob. P of damage
2. Spherical targets : Get $P(f)$, prob of damaging fr. f.
3. Others

Wethtal volume

1. Spherical (Cookie cutter) R_E

2. Ellipsoidal a, b

3. Casualty cone :

$$\phi(r) = P(b, \frac{b^2}{a^2})$$

$$\phi(r) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-t^2/2} dt$$

PV 20% cone

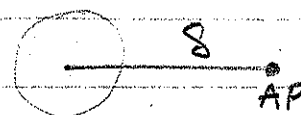
$$\phi(r) = e^{-0.272(r/a)^2}$$

Aiming Point

1. at point target

or at Center of volume target

2. off Center : S



Distribution of Burst points (normal)

1. Spherical Normal : $\frac{1}{(2\pi)^{3/2}} e^{-\frac{1}{2} \frac{r^2}{\sigma^2}}$

σ_D standard deviation $[\int r^2 g(r) dr]^{1/2}$

S_D Spherical prob error
radius of $\frac{1}{2}$ sphere

$$S_D = 1.538 \sigma_D$$

2. Ellipsoidal Normal :

$\sigma_1, \sigma_2, \sigma_3$

$$\frac{1}{(2\pi)^{3/2} \sigma_1 \sigma_2 \sigma_3} e^{-\frac{1}{2} \left[\frac{x^2}{\sigma_1^2} + \frac{y^2}{\sigma_2^2} + \frac{z^2}{\sigma_3^2} \right]}$$

(use $\sigma_1 = \sigma_2$)

Nomographs

I

$\left\{ \begin{array}{l} \text{point} \\ \text{Spherical} \end{array} \right\}$ targets with $\left\{ \begin{array}{l} \text{on} \\ \text{off} \end{array} \right\}$ center aiming

With 1. Spherical lethal volume
 2. Spherical normal dist. of burst point.

II point target with on center aiming.

With: 1. Ellipsoidal lethal volume
 2. Spherical normal dist of burst point
 or vice versa

III Spherical targets with on center aiming.

With 1. Spherical lethal volume
 2. Ellipsoidal normal dist. of burst point.

get $P(R)$ curve.

Slide Rule for

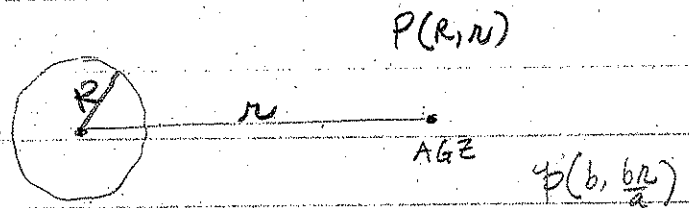
point target and $\begin{Bmatrix} \text{on} \\ \text{off} \end{Bmatrix}$ center aiming

with:

1. Spherical
2. " " " "

Casualty Curves

No Nomographs



Using spherical coverage function (analog

to Gormon's circular coverage function) for

a casualty curve we are able to quite
easily get the above results. ^(On Sph. Normal dist. case) Perhaps

in pt target case we can use other

casualty functions.

ALSO STEP FUNCTIONS